

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
18 March 2004 (18.03.2004)

PCT

(10) International Publication Number  
**WO 2004/022778 A1**

(51) International Patent Classification<sup>7</sup>: C12Q 1/68,  
G01N 33/53

(AU). O'BRIEN, Philippa [GB/AU]; 73 Ocean Street,  
Woollahra, New South Wales 20258 (AU).

(21) International Application Number:  
PCT/AU2003/001166

(74) Agent: F B RICE & CO.; 139 Rathdowne Street, Carlton,  
Victoria 3053 (AU).

(22) International Filing Date:  
5 September 2003 (05.09.2003)

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,  
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,  
MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC,  
SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA,  
UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
2002951346 5 September 2002 (05.09.2002) AU

(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,  
ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO,  
SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM,  
GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): GARVAN  
INSTITUTE OF MEDICAL RESEARCH [AU/AU]; C/-  
St Vincent's Hospital, 384 victoria Street, Darlinghurst,  
New South Wales 2010 (AU).

**Published:**

— with international search report

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): SUTHERLAND,  
Robert [NZ/AU]; 20 Northcot Road, Lindfield, New  
South Wales 2070 (AU). HENSHALL, Susan [AU/AU];  
14 Mosman Street, Mosman, New South Wales 2068

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 2004/022778 A1

(54) Title: METHODS OF DIAGNOSIS AND PROGNOSIS OF OVARIAN CANCER

(57) Abstract: The present invention provides novel genes and proteins for diagnosing ovarian cancer and/or a likelihood for survival, or recurrence of disease, wherein the expression of the genes and proteins is up-regulated or down-regulated or associated with the occurrence or recurrence of a specific scanner sub-type. The ovarian cancer-associated genes and proteins of the invention are specifically exemplified by the genes and proteins set forth in Tables 1 to 3 and the Sequence Listing.

BEST AVAILABLE COPY

## METHODS OF DIAGNOSIS AND PROGNOSIS OF OVARIAN CANCER

Field of the invention

The present invention relates to the identification of nucleic acid and protein expression profiles and nucleic acids, products, and antibodies thereto that are involved in ovarian cancer; and to the use of such expression profiles and compositions in the diagnosis, prognosis and therapy of ovarian cancer. More particularly, this invention relates to novel genes that are expressed at elevated or reduced levels in malignant tissues and uses therefor in the diagnosis of cancer or malignant tumors in human subjects. This invention also relates to the use of nucleic acid or antibody probes to specifically detect ovarian cancer cells, such as, for example, in the ovarian surface epithelium, wherein over-expression or reduced expression of nucleic acids hybridizing to the probes is highly associated with the occurrence and/or recurrence of an ovarian tumor, and/or the likelihood of patient survival. The diagnostic and prognostic test of the present invention is particularly useful for the early detection of ovarian cancer or metastases thereof, or other cancers, and for monitoring the progress of disease, such as, for example, during remission or following surgery or chemotherapy. The present invention is also directed to methods of therapy wherein the activity of a protein encoded by a diagnostic/prognostic gene described herein is modulated.

Background of the invention*1. General*

As used herein the term "derived from" shall be taken to indicate that a specified integer are obtained from a particular source albeit not necessarily directly from that source.

Unless the context requires otherwise or specifically stated to the contrary, integers, steps, or elements of the invention recited herein as singular integers, steps or elements clearly encompass both singular and plural forms of the recited integers, steps or elements.

The embodiments of the invention described herein with respect to any single embodiment shall be taken to apply *mutatis mutandis* to any other embodiment of the invention described herein.



Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated step or element or integer or group of steps or elements or integers but not the exclusion of any other step or element or integer or group of elements or integers.

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations or any two or more of said steps or features.

The present invention is not to be limited in scope by the specific examples described herein. Functionally equivalent products, compositions and methods are clearly within the scope of the invention, as described herein.

The present invention is performed without undue experimentation using, unless otherwise indicated, conventional techniques of molecular biology, microbiology, virology, recombining DNA technology, peptide synthesis in solution, solid phase peptide synthesis, and immunology. Such procedures are described, for example, in the following texts that are incorporated herein by reference:

1. Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, New York; Second Edition (1989), whole of Vols I, II, and III;
2. *DNA Cloning: A Practical Approach*, Vols. I and II (D. N. Glover, ed., 1985), IRL Press, Oxford, whole of text;
3. *Oligonucleotide Synthesis: A Practical Approach* (M. J. Gait, ed., 1984) IRL Press, Oxford, whole of text, and particularly the papers therein by Gait, pp1-22; Atkinson *et al.*, pp35-81; Sproat *et al.*, pp 83-115; and Wu *et al.*, pp 135-151;
4. *Nucleic Acid Hybridization: A Practical Approach* (B. D. Hames & S. J. Higgins, eds., 1985) IRL Press, Oxford, whole of text;
5. Perbal, B., *A Practical Guide to Molecular Cloning* (1984);
6. Wünsch, E., ed. (1974) *Synthese von Peptiden in Houben-Weyls Methoden der Organischen Chemie* (Müller, E., ed.), vol. 15, 4th edn., Parts 1 and 2, Thieme, Stuttgart.

7. Handbook of Experimental Immunology, Vols. I-IV (D. M. Weir and C. C. Blackwell, eds., 1986, Blackwell Scientific Publications).

5 This specification contains nucleotide and amino acid sequence information prepared using PatentIn Version 3.1, presented herein after the claims. Each nucleotide sequence is identified in the sequence listing by the numeric indicator <210> followed by the sequence identifier (e.g. <210>1, <210>2, <210>3, etc). The length and type of sequence (DNA, protein (PRT), etc), and source organism for each nucleotide sequence, are indicated by information provided in the numeric indicator fields <211>, <212> and  
10 <213>, respectively. Nucleotide sequences referred to in the specification are defined by the term "SEQ ID NO:", followed by the sequence identifier (eg. SEQ ID NO: 1 refers to the sequence in the sequence listing designated as <400>1).

The designation of nucleotide residues referred to herein are those recommended by the  
15 IUPAC-IUB Biochemical Nomenclature Commission, wherein A represents Adenine, C represents Cytosine, G represents Guanine, T represents thymine, Y represents a pyrimidine residue, R represents a purine residue, M represents Adenine or Cytosine, K represents Guanine or Thymine, S represents Guanine or Cytosine, W represents Adenine or Thymine, H represents a nucleotide other than Guanine, B represents a  
20 nucleotide other than Adenine, V represents a nucleotide other than Thymine, D represents a nucleotide other than Cytosine and N represents any nucleotide residue.

## 2. *Description of the related art*

Cancer is a multi-factorial disease and major cause of morbidity in humans and other  
25 animals, and deaths resulting from cancer in humans are increasing and expected to surpass deaths from heart disease in future. Carcinomas of the lung, prostate, breast, colon, pancreas, and ovary are major contributing factors to total cancer death in humans. For example, prostate cancer is the fourth most prevalent cancer and the second leading cause of cancer death in males. Similarly, cancer of the ovary is the  
30 second most common cancer of the female reproductive organs and the fourth most common cause of cancer death among females. With few exceptions, metastatic disease from carcinoma is fatal. Even if patients survive their primary cancers, recurrence or metastases are common.

35 It is widely recognized that simple and rapid tests for solid cancers or tumors have considerable clinical potential. Not only can such tests be used for the early diagnosis of

cancer but they also allow the detection of tumor recurrence following surgery and chemotherapy. A number of cancer-specific blood tests have been developed which depend upon the detection of tumor-specific antigens in the circulation (Catalona, W.J., *et al.*, 1991, "Measurement of prostate-specific antigen in serum as a screening test for prostate cancer", *N. Engl. J. Med.* 324, 1156-1161; Barrenetxea, G., *et al.*, 1998, "Use of serum tumor markers for the diagnosis and follow-up of breast cancer", *Oncology*, 55, 447-449; Cairns, P., and Sidransky, D., 1999, "Molecular methods for the diagnosis of cancer". *Biochim. Biophys. Acta.* 1423, C 11-C 18).

- 10 Ovarian cancer is the fourth most frequent cause of cancer death in females and in the United States, and accounts for approximately 13,000 deaths annually. Furthermore, ovarian cancer remains the number one killer of women with gynaecological malignant hyperplasia and the incidence is rising in industrialized countries. The etiology of the neoplastic transformation remains unknown although there is epidemiological evidence  
15 for an association with disordered endocrine function. The incidence of ovarian carcinoma is higher in nulliparous females and in those with early menopause.

Most ovarian cancers are thought to arise from the ovarian surface of epithelium (OSE). Epithelial ovarian cancer is seldom encountered in women less than 35 years of age. Its  
20 incidence increases sharply with advancing age and peaks at ages 75 to 80, with the median age being 60 years. The single most important known risk factor is a strong familial history of breast or ovarian cancer. To date, little is known about the structure and function of the OSE cells. It is known that the OSE is highly dynamic tissue that undergoes morphogenic changes, and has proliferative properties sufficient to cover the  
25 ovulatory site following ovulation. Morphological and histochemical studies suggest that the OSE has secretory, endocytotic and transport functions which are hormonally-controlled (Blaustein and Lee, *Oncol.* 8, 34-43, 1979; Nicosia and Johnson, *Int. J. Gynecol. Pathol.*, 3, 249-260, 1983; Papadaki and Beilby, *J. Cell Sci.* 8, 445-464, 1971; Anderson *et al.*, *J. Morphol.*, 150, 135-164, 1976).

30 Ovarian cancers are not readily detectable by diagnostic techniques (Siemens *et al.*, *J. Cell. Physiol.*, 134: 347-356, 1988). In fact, the diagnosis of carcinoma of the ovary is generally only possible when the disease has progressed to a late stage of development. Approximately 75% of women diagnosed with ovarian cancer are already at an advanced  
35 stage (III and IV) of the disease at their initial diagnosis. During the past 20 years, neither diagnosis nor five year survival rates have greatly improved for these patients. This is

substantially due to the high percentage of high-stage initial detection of the disease. There is therefore a need to develop new markers that improve early diagnosis and thereby reduce the percentage of high-stage initial diagnoses.

5 A number of proteinaceous ovarian tumor markers were evaluated several years ago, however these were found to be non-specific, and determined to be of low value as markers for primary ovarian cancer (Kudlacek *et al.*, *Gyn. Onc.* 35, 323-329, 1989; Rustin *et al.*, *J. Clin. Onc.*, 7, 1667-1671, 1989; Sevela *et al.*, *Am. J. Obstet. Gynecol.*, 161, 1213-1216, 1989; Omar *et al.*, *Tumor Biol.*, 10, 316-323, 1989). Several  
10 monoclonal antibodies were also shown to react with ovarian tumor associated antigens, however they were not specific for ovarian cancer and merely recognize determinants associated with high molecular weight mucin-like glycoproteins (Kenemans *et al.*, *Eur. J. Obstet. Gynecol. Reprod. Biol.* 29, 207-218, 1989; McDuffy, *Ann. Clin. Biochem.*, 26, 379-387, 1989). More recently, oncogenes associated with ovarian cancers have been  
15 identified, including *HER-2/neu (c-erbB-2)* which is over-expressed in one-third of ovarian cancers (USSN 6,075,122 by Cheever *et al.*, issued June 13, 2000), the *fms* oncogene, and abnormalities in the *p53* gene, which are seen in about half of ovarian cancers.

Whilst previously identified markers for carcinomas of the ovary have facilitated efforts to  
20 diagnose and treat these serious diseases, there is a clear need for the identification of additional markers and therapeutic targets. The identification of tumor markers that are amenable to the early-stage detection of localized tumors is critical for more effective management of carcinomas of the ovary.

## 25 Summary of the invention

In work leading up to the present invention, the inventors sought to identify nucleic acid markers that were diagnostic of ovarian cancers generally, or diagnostic of specific ovarian cancers such as, for example, serous ovarian cancer (SOC), mucinous ovarian cancer (MOC), non-invasive (borderline ovarian cancer or low malignant potential  
30 ovarian cancer), mixed phenotype ovarian cancer, endometrioid ovarian cancer (EnOC) and clear cell ovarian cancer (CICA), papillary serous ovarian cancer, Brenner cell or undifferentiated adenocarcinoma, by virtue of their modulated expression in cancer tissues derived from a patient cohort compared to their expression in healthy or non-cancerous cells and tissues. Additionally, the inventors sought to determine whether any  
35 correlation exists between the expression of any particular gene in a subject having ovarian cancer and the survival, or likelihood for survival, of the subject during the

medium to long term (i.e. in the period between about 1-2 years from primary diagnosis, or longer). The inventors also sought to determine whether any correlation exists between the expression of any particular gene in a subject following treatment for ovarian cancer and the recurrence, or likelihood for recurrence, of ovarian cancer in the subject during the medium to long term (i.e. in the period between about 1-2 years from primary diagnosis, or longer).

As exemplified herein, the inventors identified a number of genes whose expression is altered (up-regulated or down-regulated) in individuals with ovarian cancer compared to healthy individuals., eg., subjects who do not have ovarian cancer. The particular genes are identified in Tables 1 and 2. Preferably, the genes are selected from the group of candidate genes set forth in Table 3.

The list of genes and proteins exemplified herein by Table 1 were identified by a statistical analysis as outlined in the examples which gave a P-value, eg., by comparison of expression to the expression of that gene in normal ovaries.

Accordingly, one aspect of the present invention provides a method of detecting an ovarian cancer-associated transcript in a biological sample, the method comprising contacting the biological sample with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Table 1 or 2 or 3. Preferably the percentage identity to a sequence disclosed in any one of Tables 1-3 is at least about 85% or 90% or 95%, and still more preferably at least about 98% or 99%.

In a preferred embodiment, the present invention provides a method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein a modified level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:

- (i) a sequence comprising at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 46, 48, 50, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;

- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 46, 48, 50, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- (iii) a sequence that is at least about 80% identical to a sequence selected from the group consisting of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 46, 48, 50, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- (iv) a sequence that encodes an amino acid sequence selected from the group consisting of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 47, 49, 51, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84; and
- (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

In a preferred embodiment, the present invention provides a method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein a modified level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:

- (i) a sequence comprising at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 46, 48, 50, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 46, 48, 50, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- (iii) a sequence that is at least about 80% identical to a sequence selected from the group consisting of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27,

29, 31, 33, 35, 37, 39, 41, 43, 45, 46, 48, 50, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;

- (iv) a sequence that encodes an amino acid sequence selected from the group consisting of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 47, 49, 51, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84; and
- (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

10 Even more preferably, the present invention provides a method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein a modified level of hybridization of the probe for the subject being tested  
15 compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:

- (i) a sequence comprising at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1, 5, 7, 9, 11, 13, 15, 17, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 45, 46, 48, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;

- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1, 5, 7, 9, 11, 13, 15, 17, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 45, 46, 48, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;

- (iii) a sequence that is at least about 80% identical to a sequence selected from the group consisting of SEQ ID NOs: 1, 5, 7, 9, 11, 13, 15, 17, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 45, 46, 48, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;

- (iv) a sequence that encodes an amino acid sequence selected from the group consisting of SEQ ID NOs: 2, 6, 8, 10, 12, 14, 16, 18, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 47, 49, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84; and

- 35 (v) a sequence that is complementary to (i) or (ii) or (iii) or (iv).

As used herein, the term "modified level" includes an enhanced, increased or elevated level of an integer being assayed, or alternatively, a reduced or decreased level of an integer being assayed.

- 5 In one embodiment an elevated, enhanced or increased level of expression of the nucleic acid is detected. In accordance with this embodiment, the present invention provides a method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and
- 10 then detecting the hybridization wherein an enhanced level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian ovarian cancer, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:
- 15 (i) a sequence comprising at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 or 2 other than a nucleic acid having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200;
- 20 (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 or 2 other than a nucleic acid having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200;
- (iii) a sequence that is at least about 80% identical to (i) or (ii);
- 25 (iv) a sequence that encodes a polypeptide encoded by a nucleic acid set forth in Table 1 or 2 other than a nucleic acid having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200; and
- (v) a sequence that is complementary to any one of the sequences set forth in (i) or
- 30 (ii) or (iii) or (iv).

In a preferred embodiment, the present invention provides a method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe

35 for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein an enhanced level of hybridization of the probe for the subject



being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:

- 5 (i) a sequence comprising at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 7, 9, 11, 13, 15, 17, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 45, 46, 48, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions  
10 to at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 7, 9, 11, 13, 15, 17, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 45, 46, 48, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- (iii) a sequence that is at least about 80% identical to a sequence selected from the  
15 group consisting of SEQ ID NOs: 7, 9, 11, 13, 15, 17, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 45, 46, 48, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- (iv) a sequence that encodes an amino acid sequence selected from the group  
20 consisting of SEQ ID NOs: 8, 10, 12, 14, 16, 18, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 47, 49, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84;  
and
- (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

- 25 In an alternative preferred embodiment, a reduced level of a diagnostic marker is indicative of ovarian cancer. In accordance with this embodiment, the present invention provides a method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for  
30 hybridization to occur and then detecting the hybridization wherein a reduced level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian ovarian cancer, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:

- (i) a sequence comprising at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200;
  - (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200;
  - (iii) a sequence that is at least about 80% identical to (i) or (ii);
  - (iv) a sequence that encodes a polypeptide encoded by a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200; and
  - (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).
- 15 In a preferred embodiment, the present invention provides a method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein a reduced level of hybridization of the probe for the subject being
- 20 tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:
- (i) a sequence comprising at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1, 3, and 5;
  - (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1, 3, and 5;
  - (iii) a sequence that is at least about 80% identical to a sequence selected from the group consisting of SEQ ID NOs: 1, 3, and 5;
  - (iv) a sequence that encodes an amino acid sequence selected from the group consisting of SEQ ID NOs: 2, 4, and 6; and
  - (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).
- 35 Preferably, the ovarian cancer that is diagnosed according to the present invention is an epithelial ovarian cancer, such as, for example, serous ovarian cancer, non-invasive

ovarian cancer, mixed phenotype ovarian cancer, mucinous ovarian cancer, endometrioid ovarian cancer, clear cell ovarian cancer, papillary serous ovarian cancer, Brenner cell or undifferentiated adenocarcinoma. As will be apparent from the preferred embodiments described below, certain of the genes represented in Table 1, Table 2 and Table 3 are expressed at modified levels in subjects having serous or mucinous ovarian cancers. Data presented in Figures 1-4 also exemplify novel diagnostics and prognostics for serous ovarian cancer, mucinous ovarian cancer, endometrioid ovarian cancer and clear cell ovarian cancer.

As exemplified herein by Table 2, the present inventors have identified those genes having an elevated or reduced average ratio of expression of specific genes between ovarian cancer patients vs non-ovarian cancer patients, wherein a high ratio in Table 2 indicates an enhanced expression in an ovarian cancer patients and wherein a negative ratio indicates that a reduced expression in an ovarian cancer patient.

15

In an alternative preferred embodiment, the present invention provides a method of diagnosing a serous ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein a modified level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has a serous ovarian cancer, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:

- 25 (i) a sequence comprising at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 2 or as set forth in Table 1 and having an Accession Number selected from the group consisting of: U62801, D49441, X51630, and AB018305;
- 30 (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 2 or as set forth in Table 1 and having an Accession Number selected from the group consisting of: U62801, D49441, X51630, And AB018305;
- (iii) a sequence that is at least about 80% identical to (i) or (ii);
- 35 (iv) a sequence that encodes a polypeptide encoded by a nucleic acid set forth in Table 2 or as set forth in Table 1 and having an Accession Number selected from the group consisting of: U62801, D49441, X51630, And AB018305; and

- (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

In a further alternative preferred embodiment, the present invention provides a method of  
5 diagnosing a mucinous ovarian cancer in a human or animal subject being tested said  
method comprising contacting a biological sample from said subject being tested with a  
nucleic acid probe for a time and under conditions sufficient for hybridization to occur and  
then detecting the hybridization wherein an elevated level of hybridization of the probe for  
the subject being tested compared to the hybridization obtained for a control subject not  
10 having ovarian cancer indicates that the subject being tested has a mucinous ovarian  
cancer, and wherein said nucleic acid probe comprises a sequence selected from the  
group consisting of:

- (i) a sequence comprising at least about 20 contiguous nucleotides from a nucleic  
acid set forth in Table 1 and having an Accession Number selected from the group  
15 consisting of: NM\_006149, AA315933, U47732, NM\_005588, AW503395,  
NM\_004063, AI073913, AI928445, NM\_022454, W40460, AA132961 and  
AF111856;
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions  
to at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1  
20 and having an Accession Number selected from the group consisting of:  
NM\_006149, AA315933, U47732, NM\_005588, AW503395, NM\_004063,  
AI073913, AI928445, NM\_022454, W40460, AA132961 and AF111856;
- (iii) a sequence that is at least about 80% identical to (i) or (ii);
- (iv) a sequence that encodes a polypeptide encoded by a nucleic acid set forth in  
25 Table 1 and having an Accession Number selected from the group consisting of:  
NM\_006149, AA315933, U47732, NM\_005588, AW503395, NM\_004063,  
AI073913, AI928445, NM\_022454, W40460, AA132961 and AF111856; and
- (v) a sequence that is complementary to any one of the sequences set forth in (i) or  
(ii) or (iii) or (iv).

30

In a preferred embodiment, the present invention provides a method of diagnosing a  
mucinous ovarian cancer in a human or animal subject being tested said method  
comprising contacting a biological sample from said subject being tested with a nucleic  
acid probe for a time and under conditions sufficient for hybridization to occur and then  
35 detecting the hybridization wherein an enhanced level of hybridization of the probe for  
the subject being tested compared to the hybridization obtained for a control subject not

having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:

- (i) a sequence comprising at least about 20 contiguous nucleotides from SEQ ID NO: 57 or 59 or 61;
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from SEQ ID NO: 57 or 59 or 61;
- (iii) a sequence that is at least about 80% identical to SEQ ID NO: 57 or 59 or 61;
- (iv) a sequence that encodes the amino acid sequence set forth in SEQ ID NO: 58 or 60 or 62; and
- (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

Those skilled in the art will be aware that as a carcinoma progresses, metastases occur in organs and tissues outside the site of the primary tumor. For example, in the case of ovarian cancer, metastases commonly appear in a tissue selected from the group consisting of omentum, abdominal fluid, lymph nodes, lung, liver, brain, and bone. Accordingly, the term "ovarian cancer" as used herein shall be taken to include an early or developed tumor of the ovary, such as, for example, any one or more of a number of cancers of epithelial origin, such as serous, mucinous, endometrioid, clear cell, papillary serous, Brenner cell or undifferentiated adenocarcinoma, non-invasive ovarian cancer such as borderline ovarian cancer or low-malignant potential ovarian cancer, or a mixed phenotype ovarian cancer, and optionally, any metastases outside the ovary that occurs in a subject having a primary tumor of the ovary.

As used herein, the term "diagnosis", and variants thereof, such as, but not limited to "diagnose", "diagnosed" or "diagnosing" shall not be limited to a primary diagnosis of a clinical state, however should be taken to include any primary diagnosis or prognosis of a clinical state. For example, the "diagnostic assay" formats described herein are equally relevant to assessing the remission of a patient, or monitoring disease recurrence, or tumor recurrence, such as following surgery or chemotherapy, or determining the appearance of metastases of a primary tumor. All such uses of the assays described herein are encompassed by the present invention.

Both classical hybridization and amplification formats, and combinations thereof, are encompassed by the invention. In one embodiment, the hybridization comprises

performing a nucleic acid hybridization reaction between a labeled probe and a second nucleic acid in the biological sample from the subject being tested, and detecting the label. In another embodiment, the hybridization comprising performing a nucleic acid amplification reaction eg., polymerase chain reaction (PCR), wherein the probe consists of a nucleic acid primer and nucleic acid copies of the nucleic acid in the biological sample are amplified. As will be known to the skilled artisan, amplification may proceed classical nucleic acid hybridization detection systems, to enhance specificity of detection, particularly in the case of less abundant mRNA species in the sample.

- 10 In a preferred embodiment, the polynucleotide is immobilised on a solid surface.

The present invention clearly encompasses nucleic acid-based methods and protein-based methods for diagnosing cancer in humans and other mammals.

- 15 Accordingly, in a related embodiment, the present invention provides a method of detecting an ovarian cancer-associated polypeptide in a biological sample the method comprising contacting the biological sample with an antibody that binds specifically to an ovarian cancer-associated polypeptide in the biological sample, the polypeptide being encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-3.

Preferably the percentage identity to a sequence disclosed in any one of Tables 1-3 is at least about 85% or 90% or 95%, and still more preferably at least about 98% or 99%.

- 25 In a preferred embodiment, the present invention provides a method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein a modified level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a sequence having at least about 80% identity to a sequence selected from the group consisting of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34,

36, 38, 40, 42, 44, 47, 49, 51, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84.

In a preferred embodiment, the present invention provides a method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein a modified level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a sequence having at least about 80% identity to a sequence selected from the group consisting of SEQ ID NOs: 2, 6, 8, 10, 12, 14, 16, 18, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 47, 49, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84.

In one embodiment an elevated, enhanced or increased level of expression of the antigen-antibody complex is detected. In accordance with this embodiment, the present invention provides a method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein an enhanced level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a polypeptide encoded by a nucleic acid set forth in Table 1 or 2 other than a nucleic acid having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200.

In a preferred embodiment, the present invention provides a method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein an enhanced level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for

a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a sequence having at least about 80% identity to a sequence selected from the group consisting of SEQ ID NOs: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 47, 49, 51, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84.

In an alternative preferred embodiment, a reduced level of a diagnostic marker is indicative of ovarian cancer. In accordance with this embodiment, the present invention provides a method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein a reduced level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a polypeptide encoded by a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200.

In a preferred embodiment, the present invention provides a method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein a reduced level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a sequence having at least about 80% identity to a sequence selected from the group consisting of SEQ ID NOs: 2, 4, and 6.

Preferably, the ovarian cancer that is diagnosed according to the present invention is an epithelial ovarian cancer, such as, for example, serous ovarian cancer or mucinous ovarian cancer.



In an alternative preferred embodiment, the present invention provides a method of diagnosing a serous ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein a modified level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has a serous ovarian cancer, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a polypeptide encoded by a nucleic acid set forth in Table 2 or as set forth in Table 1 and having an Accession Number selected from the group consisting of: U62801, D49441, X51630, And AB018305.

In a further alternative preferred embodiment, the present invention provides a method of diagnosing a mucinous ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein a reduced level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has a mucinous ovarian cancer, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a polypeptide encoded by a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: NM\_006149, AA315933, U47732, NM\_005588, AW503395, NM\_004063, AI073913, AI928445, NM\_022454, W40460, AA132961 and AF111856.

In a preferred embodiment, the present invention provides a method of diagnosing a mucinous ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein an enhanced level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has a mucinous ovarian cancer, and wherein said antibody binds to a

polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a sequence having at least about 80% identity to SEQ ID NO: 58 or 60 or 62.

- 5 In a further related embodiment, the present invention provides a method of detecting an ovarian cancer-associated antibody in a biological sample the method comprising contacting the biological sample with a polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-3, wherein the polypeptide specifically binds to the ovarian cancer-associated  
10 antibody.

Preferably, in the above methods, the biological sample is contacted with a plurality of the polynucleotides, polypeptides or antibodies referred to above.

- 15 In a particularly preferred embodiment, the present invention provides an antibody-based multiplex assay for determining the likelihood of survival of a subject from an ovarian cancer. In one embodiment, the invention provides a method of determining the likelihood of survival of a subject suffering from a serous ovarian cancer, said method comprising contacting a biological sample from said subject being tested with at least two  
20 antibodies for a time and under conditions sufficient for antigen-antibody complexes to form and then detecting the complexes wherein an enhanced level of the antigen-antibody complexes for the subject being tested compared to the amount of the antigen-antibody complexes formed for a control subject not having ovarian cancer indicates that the subject being tested has a poor probability of survival, and wherein one antibody  
25 binds to an sFRP polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 72 and wherein one antibody binds to a SOCS3 polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 74.

- The present invention is not to be limited by the source or nature of the biological  
30 sample. In one embodiment, the biological sample is from a patient undergoing a therapeutic regimen to treat ovarian cancer. In an alternative preferred embodiment, the biological sample is from a patient suspected of having ovarian cancer.

- In addition to providing up-regulated and down-regulated genes, the list of genes and  
35 proteins exemplified herein by Table 1 were identified by a statistical analysis as outlined in the examples which gave a P-value, eg., by comparison of expression to

clinicopathological parameters for disease recurrence or patient survival. Accordingly, the present invention is particularly useful for prognostic applications, in particular for assessing the medium-to-long term survival of a subject having an ovarian cancer, or alternatively or in addition, for assessing the likelihood of disease recurrence.

5

Accordingly, a further aspect of the present invention provides a method of monitoring the efficacy of a therapeutic treatment of ovarian cancer, the method comprising:

- (i) providing a biological sample from a patient undergoing the therapeutic treatment; and
- 10 (ii) determining the level of a ovarian cancer-associated transcript in the biological sample by contacting the biological sample with a polynucleotide that selectively hybridizes to a sequence having at least about 80% identity to a sequence as shown in any one of Tables 1-3, thereby monitoring the efficacy of the therapy.

15

Preferably the method further comprises comparing the level of the ovarian cancer-associated transcript to a level of the ovarian cancer-associated transcript in a biological sample from the patient prior to, or earlier in, the therapeutic treatment.

20

In a related embodiment, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of ovarian cancer, the method comprising :

- (i) providing a biological sample from a patient undergoing the therapeutic treatment; and
- 25 (ii) determining the level of a ovarian cancer-associated antibody in the biological sample by contacting the biological sample with a polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-3, wherein the polypeptide specifically binds to the ovarian cancer-associated antibody, thereby monitoring the efficacy of the therapy.

30

Preferably the method further comprises comparing the level of the ovarian cancer-associated antibody to a level of the ovarian cancer-associated antibody in a biological sample from the patient prior to, or earlier in, the therapeutic treatment.

In a further related embodiment, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of ovarian cancer, the method comprising :

- (i) providing a biological sample from a patient undergoing the therapeutic treatment; and
- 5 (ii) determining the level of a ovarian cancer-associated polypeptide in the biological sample by contacting the biological sample with an antibody, wherein the antibody specifically binds to a polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-3, thereby monitoring the efficacy of the therapy.

10

Preferably the method further comprises comparing the level of the ovarian cancer-associated polypeptide to a level of the ovarian cancer-associated polypeptide in a biological sample from the patient prior to, or earlier in, the therapeutic treatment.

- 15 It will also be apparent from the following preferred embodiments, that the expression of certain genes listed in Table 1 and Table 3 is statistically correlated with survival and death of patients having ovarian cancer, wherein a low P value indicates an enhanced likelihood that a patient having altered expression of the gene will die from the cancer.

- 20 Accordingly, in one embodiment, the present invention provides a method of determining the likelihood of survival of a subject suffering from an ovarian cancer, said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein an elevated level of hybridization of the probe for the
- 25 subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has a poor probability of survival, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:

- (i) a sequence comprising at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: NM\_003014, AA046217, NM\_015902, T83882, AB040888, AA628980, AI623351, AW614420, AA243499, AF251237, AI970797, AF145713, X78565, T97307, BE243845, AW068302, AL133561, BE313555, X07820, AI973016, AF084545, U41518, Z11894, AW138190, BE086548, W47196, AI796870, X02761, AW968613, AW972565, AF045229, AW953853, U52426, F06700, AI798863, H52761, BE546947, AU076643, U20536, AA581602,
- 30
- 35

- AJ245210, X65965, AI806770, BE386490, AW581992, U77534, AL034417, L10343, AW518944, W28729, AI640160, U11862, AW295980, X59135, BE466173, AI354722, M90464, AA829286, AI333771, BE465867, NM\_014992, BE616902, AA430373, R27430, BE387335, AW264102, AW952323, AA088177, BE614567, AL079658, NM\_002776, BE261944, NM\_006379, AI002238, X81789, NM\_002122, AB001914, AA311919, AI381750, AA292998, BE439580, AI677897, N72403, BE003054, AL035588, AI080491, AW770994, H24177, AF146761, NM\_001955, AI680737, AI752666, AA505445, BE246649, and NM\_003955;
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: NM\_003014, AA046217, NM\_015902, T83882, AB040888, AA628980, AI623351, AW614420, AA243499, AF251237, AI970797, AF145713, X78565, T97307, BE243845, AW068302, AL133561, BE313555, X07820, AI973016, AF084545, U41518, Z11894, AW138190, BE086548, W47196, AI796870, X02761, AW968613, AW972565, AF045229, AW953853, U52426, F06700, AI798863, H52761, BE546947, AU076643, U20536, AA581602, AJ245210, X65965, AI806770, BE386490, AW581992, U77534, AL034417, L10343, AW518944, W28729, AI640160, U11862, AW295980, X59135, BE466173, AI354722, M90464, AA829286, AI333771, BE465867, NM\_014992, BE616902, AA430373, R27430, BE387335, AW264102, AW952323, AA088177, BE614567, AL079658, NM\_002776, BE261944, NM\_006379, AI002238, X81789, NM\_002122, AB001914, AA311919, AI381750, AA292998, BE439580, AI677897, N72403, BE003054, AL035588, AI080491, AW770994, H24177, AF146761, NM\_001955, AI680737, AI752666, AA505445, BE246649, and NM\_003955;
- (iii) a sequence that is at least about 80% identical to (i) or (ii);
- (iv) a sequence that encodes a polypeptide encoded by a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: NM\_003014, AA046217, NM\_015902, T83882, AB040888, AA628980, AI623351, AW614420, AA243499, AF251237, AI970797, AF145713, X78565, T97307, BE243845, AW068302, AL133561, BE313555, X07820, AI973016, AF084545, U41518, Z11894, AW138190, BE086548, W47196, AI796870, X02761, AW968613, AW972565, AF045229, AW953853, U52426, F06700, AI798863, H52761, BE546947, AU076643, U20536, AA581602, AJ245210, X65965, AI806770, BE386490, AW581992, U77534, AL034417, L10343, AW518944, W28729, AI640160, U11862, AW295980, X59135, BE466173, AI354722,

M90464, AA829286, AI333771, BE465867, NM\_014992, BE616902, AA430373, R27430, BE387335, AW264102, AW952323, AA088177, BE614567, AL079658, NM\_002776, BE261944, NM\_006379, AI002238, X81789, NM\_002122, AB001914, AA311919, AI381750, AA292998, BE439580, AI677897, N72403, BE003054, AL035588, AI080491, AW770994, H24177, AF146761, NM\_001955, AI680737, AI752666, AA505445, BE246649, and NM\_003955; and

(v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

10 In a preferred embodiment, the present invention provides a method of determining the likelihood of survival of a subject suffering from an ovarian cancer, said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein an elevated level of hybridization of the probe for the  
15 subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has a poor probability of survival, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:

- (i) a sequence comprising at least about 20 contiguous nucleotides from a sequence  
20 selected from the group consisting of SEQ ID NOs: 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- 25 (iii) a sequence that is at least about 80% identical to a sequence selected from the group consisting of SEQ ID NOs: 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- (iv) a sequence that encodes an amino acid sequence selected from the group consisting of SEQ ID NOs: 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84; and
- (v) a sequence that is complementary to (i) or (ii) or (iii) or (iv).

30

In an alternative preferred embodiment, the present invention provides a method of determining the likelihood of survival of a subject suffering from an ovarian cancer, said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to  
35 form and then detecting the complex wherein an enhanced level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody

complex formed for a control subject not having ovarian cancer indicates that the subject being tested has a poor probability of survival, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a sequence encoded by a nucleic acid set forth in Table 1 and

5 having an Accession Number selected from the group consisting of: NM\_003014, AA046217, NM\_015902, T83882, AB040888, AA628980, AI623351, AW614420, AA243499, AF251237, AI970797, AF145713, X78565, T97307, BE243845, AW068302, AL133561, BE313555, X07820, AI973016, AF084545, U41518, Z11894, AW138190, BE086548, W47196, AI796870, X02761, AW968613, AW972565, AF045229,

10 AW953853, U52426, F06700, AI798863, H52761, BE546947, AU076643, U20536, AA581602, AJ245210, X65965, AI806770, BE386490, AW581992, U77534, AL034417, L10343, AW518944, W28729, AI640160, U11862, AW295980, X59135, BE466173, AI354722, M90464, AA829286, AI333771, BE465867, NM\_014992, BE616902, AA430373, R27430, BE387335, AW264102, AW952323, AA088177, BE614567,

15 AL079658, NM\_002776, BE261944, NM\_006379, AI002238, X81789, NM\_002122, AB001914, AA311919, AI381750, AA292998, BE439580, AI677897, N72403, BE003054, AL035588, AI080491, AW770994, H24177, AF146761, NM\_001955, AI680737, AI752666, AA505445, BE246649, and NM\_003955.

20 In an alternative preferred embodiment, the present invention provides a method of determining the likelihood of survival of a subject suffering from an ovarian cancer, said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein an enhanced level of the antigen-antibody

25 complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has a poor probability of survival, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a sequence having at least about 80% identity to a sequence

30 selected from the group consisting of SEQ ID NOs: 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84.

In a particularly preferred embodiment, the present invention provides a marker for determining the likelihood of a subject surviving from serous cancer. In accordance with

35 this embodiment of the invention, there is provided a method of determining the likelihood of survival of a subject suffering from a serous ovarian cancer, said method

comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein an elevated level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has a poor probability of survival, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:

- (i) a sequence comprising at least about 20 contiguous nucleotides from a nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO: 71 or 73;
- 10 (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO: 71 or 73;
- (iii) a sequence that is at least about 80% identical to (i) or (ii) and encoding an sFRP protein or a SOCS3 protein;
- 15 (iv) a sequence that encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 72 or 74; and
- (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

20 In an alternative preferred embodiment, the present invention provides a method of determining the likelihood of survival of a subject suffering from a serous ovarian cancer, said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein an enhanced level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has a poor probability of survival, and wherein said antibody binds to an sFRP polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 72 or a SOCS3 polypeptide comprising the amino acid sequence set forth in  
25  
30 SEQ ID NO: 74 or.

It will also be apparent from the following preferred embodiments, that the expression of certain genes listed in Table 1 and Table 3 is statistically correlated with recurrence of ovarian cancer, wherein a low P value indicates an enhanced likelihood that a patient  
35 having altered expression of the gene will experience recurrence of the disease.



- In yet another preferred embodiment, the present invention provides a method of determining the likelihood that a subject will suffer from a recurrence of an ovarian cancer, said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein an elevated level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has a high probability of recurrence, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:
- 5 (i) a sequence comprising at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: M86849, AW963419, BE298665, AK000637, BE077546, T97307, R24601, BE090176, AA393907, W28729, BE313754, AW673081, AA356694, L08239, BE397649, NM\_012317, NM\_000947, AJ250562, AL040183, BE207573, BE564162, BE439580, AW067800, AA569756, AW138190, AF126245, L10343, NM\_002514, AI863735, NM\_005397, W26391, H15474, U51166, AA243499, AW408807, AI738719, AB040888, BE313077, AI677897, C14898, AI821730, AF007393, H65423, N46243, AA095971, U20350, NM\_005756, D19589, AW957446, AW294647, BE159718, AI888490, AA022569, BE147740, AI798863, BE464341, AL080235, AI557212, X75208, AA628980, BE242587, NM\_005512, AW953853, AU076611, AW968613, AL353944, BE614149, AA292998, H12912, AA188763, AK000596, AI970797, AW519204, Z42387, AF145713, AA972412, AK001564, AW959861, BE313555, W25005, AI193356, AF111106, AI130740, AA985190, BE221880, AF084545, R26584, AW247380, AA364261, U25849, AF262992, AW342140, AL133572, AI497778, AI745379, U51712, AW375974, AF251237, NM\_000636, AA130986, AA216363, AA628980, AA811657, AA897108, AB040888, AF212225, AI089575, AI282028, AI368826, AI718702, AI827248, AK002039, AL109791, AW090198, AW296454, AW445034, AW452948, AW470411, AW885727, AW970859, AW979189, BE165866, BE175582, BE242587, BE271927, BE439580, BE464016, D63216, F34856, M83822, N33937, N49068, N51357, N80486, NM\_000954, NM\_005756, NM\_016652, R26584, R31178, W05391, W25005, W45393, W68815, X65965, X76732 and Z45051,
- 20 (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: M86849,
- 35

- AW963419, BE298665, AK000637, BE077546, T97307, R24601, BE090176, AA393907, W28729, BE313754, AW673081, AA356694, L08239, BE397649, NM\_012317, NM\_000947, AJ250562, AL040183, BE207573, BE564162, BE439580, AW067800, AA569756, AW138190, AF126245, L10343, NM\_002514,
- 5 AI863735, NM\_005397, W26391, H15474, U51166, AA243499, AW408807, AI738719, AB040888, BE313077, AI677897, C14898, AI821730, AF007393, H65423, N46243, AA095971, U20350, NM\_005756, D19589, AW957446, AW294647, BE159718, AI888490, AA022569, BE147740, AI798863, BE464341, AL080235, AI557212, X75208, AA628980, BE242587, NM\_005512, AW953853,
- 10 AU076611, AW968613, AL353944, BE614149, AA292998, H12912, AA188763, AK000596, AI970797, AW519204, Z42387, AF145713, AA972412, AK001564, AW959861, BE313555, W25005, AI193356, AF111106, AI130740, AA985190, BE221880, AF084545, R26584, AW247380, AA364261, U25849, AF262992, AW342140, AL133572, AI497778, AI745379, U51712, AW375974, AF251237,
- 15 NM\_000636, AA130986, AA216363, AA628980, AA811657, AA897108, AB040888, AF212225, AI089575, AI282028, AI368826, AI718702, AI827248, AK002039, AL109791, AW090198, AW296454, AW445034, AW452948, AW470411, AW885727, AW970859, AW979189, BE165866, BE175582, BE242587, BE271927, BE439580, BE464016, D63216, F34856, M83822,
- 20 N33937, N49068, N51357, N80486, NM\_000954, NM\_005756, NM\_016652, R26584, R31178, W05391, W25005, W45393, W68815, X65965, X76732 and Z45051;
- (iii) a sequence that is at least about 80% identical to (i) or (ii);
- (iv) a sequence that encodes a polypeptide encoded by a nucleic acid set forth in
- 25 Table 1 and having an Accession Number selected from the group consisting of: M86849, AW963419, BE298665, AK000637, BE077546, T97307, R24601, BE090176, AA393907, W28729, BE313754, AW673081, AA356694, L08239, BE397649, NM\_012317, NM\_000947, AJ250562, AL040183, BE207573, BE564162, BE439580, AW067800, AA569756, AW138190, AF126245, L10343,
- 30 NM\_002514, AI863735, NM\_005397, W26391, H15474, U51166, AA243499, AW408807, AI738719, AB040888, BE313077, AI677897, C14898, AI821730, AF007393, H65423, N46243, AA095971, U20350, NM\_005756, D19589, AW957446, AW294647, BE159718, AI888490, AA022569, BE147740, AI798863, BE464341, AL080235, AI557212, X75208, AA628980, BE242587, NM\_005512,
- 35 AW953853, AU076611, AW968613, AL353944, BE614149, AA292998, H12912, AA188763, AK000596, AI970797, AW519204, Z42387, AF145713, AA972412,

AK001564, AW959861, BE313555, W25005, AI193356, AF111106, AI130740, AA985190, BE221880, AF084545, R26584, AW247380, AA364261, U25849, AF262992, AW342140, AL133572, AI497778, AI745379, U51712, AW375974, AF251237, NM\_000636, AA130986, AA216363, AA628980, AA811657, 5 AA897108, AB040888, AF212225, AI089575, AI282028, AI368826, AI718702, AI827248, AK002039, AL109791, AW090198, AW296454, AW445034, AW452948, AW470411, AW885727, AW970859, AW979189, BE165866, BE175582, BE242587, BE271927, BE439580, BE464016, D63216, F34856, M83822, N33937, N49068, N51357, N80486, NM\_000954, NM\_005756, 10 NM\_016652, R26584, R31178, W05391, W25005, W45393, W68815, X65965, X76732 and Z45051; and

(v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

15 In an alternative preferred embodiment, the present invention provides a method of determining the likelihood that a subject will suffer from a recurrence of an ovarian cancer, said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein an enhanced level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has a high probability of recurrence, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a sequence encoded by a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: M86849, AW963419, BE298665, AK000637, BE077546, T97307, R24601, BE090176, AA393907, W28729, BE313754, AW673081, AA356694, L08239, BE397649, NM\_012317, NM\_000947, AJ250562, AL040183, BE207573, BE564162, BE439580, AW067800, AA569756, AW138190, AF126245, L10343, NM\_002514, 25 AI863735, NM\_005397, W26391, H15474, U51166, AA243499, AW408807, AI738719, AB040888, BE313077, AI677897, C14898, AI821730, AF007393, H65423, N46243, AA095971, U20350, NM\_005756, D19589, AW957446, AW294647, BE159718, AI888490, AA022569, BE147740, AI798863, BE464341, AL080235, AI557212, X75208, AA628980, BE242587, NM\_005512, AW953853, AU076611, AW968613, AL353944, 30 BE614149, AA292998, H12912, AA188763, AK000596, AI970797, AW519204, Z42387, AF145713, AA972412, AK001564, AW959861, BE313555, W25005, AI193356,

AF111106, AI130740, AA985190, BE221880, AF084545, R26584, AW247380, AA364261, U25849, AF262992, AW342140, AL133572, AI497778, AI745379, U51712, AW375974, AF251237, NM\_000636, AA130986, AA216363, AA628980, AA811657, AA897108, AB040888, AF212225, AI089575, AI282028, AI368826, AI718702, 5 AI827248, AK002039, AL109791, AW090198, AW296454, AW445034, AW452948, AW470411, AW885727, AW970859, AW979189, BE165866, BE175582, BE242587, BE271927, BE439580, BE464016, D63216, F34856, M83822, N33937, N49068, N51357, N80486, NM\_000954, NM\_005756, NM\_016652, R26584, R31178, W05391, W25005, W45393, W68815, X65965, X76732 and Z45051.

10

The recurrence of ovarian cancer is a clinical recurrence as determined by the presence of one or more clinical symptoms of an ovarian cancer, such as, for example, a metastases, or alternatively, as determined in a biochemical test, immunological test or serological test such as, for example, a cross-reactivity in a biological sample to a CA125 15 antibody.

Preferably, the recurrence is capable of being detected at least about 2 years from treatment, more preferably about 2-3 years from treatment, and even more preferably about 4 or 5 or 10 years from treatment.

20

Preferably, in the above diagnostic and/or prognostic methods, the biological sample is contacted with a plurality of the nucleic acids and/or polypeptides and/or antibodies referred to above. In a particularly preferred embodiment, multiplex assays are performed to detect enhanced expression at least of sFRP4 and SOC3 at the protein 25 level (eg., using antigen-based or antibody-based assays) or at the mRNA level (eg., by detecting elevated levels of mRNA transcripts).

A further embodiment of the present invention provides a method of diagnosing epithelial ovarian cancer by detecting aberrant methylation of a promoter that regulates expression 30 of a tumor suppressor gene eg., MCC. In particular, the present invention contemplates the detection of hypermethylation of the promoter of a tumor suppressor gene. Without being bound by any theory or mode of action, such hypermethylation leads to gene inactivation, thereby reducing expression of the tumor suppressor gene and permitting oncogenesis. In one preferred embodiment, the present invention provides a method of 35 diagnosing an ovarian cancer in a human or animal subject being tested said method comprising determining aberrant methylation in a promoter sequence that regulates

expression of a tumor suppressor gene in a biological sample from said subject compared to the methylation of the promoter in nucleic acid obtained for a control subject not having ovarian cancer wherein said aberrant methylation indicates that the subject being tested has an ovarian ovarian cancer.

5

In a further aspect, the present invention provides a method for identifying a compound that modulates an ovarian cancer-associated polypeptide, the method comprising :

- (i) contacting the compound with a ovarian cancer-associated polypeptide, the polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-3; and
- (ii) determining the functional effect of the compound upon the polypeptide.

10

The functional effect may, for example, be a physical effect or a chemical effect. In one embodiment, the functional effect is determined by measuring ligand binding to the polypeptide. In a particular embodiment, the polypeptide is expressed in a eukaryotic host cell or cell membrane. Preferably the polypeptide is recombinant.

15

In another aspect, the present invention provides a method of inhibiting proliferation of a ovarian tumour cell, which method comprises contacting said cell with a compound identified using the method *supra* for identifying a compound that modulates an ovarian cancer-associated polypeptide.

20

In a further aspect, the present invention provides a method of inhibiting proliferation of a ovarian cancer-associated cell to treat ovarian cancer in a patient, the method comprising the step of administering to the patient a therapeutically effective amount of a compound identified using the method *supra* for identifying a compound that modulates an ovarian cancer-associated polypeptide.

25

In a further aspect, the present invention provides a drug screening assay comprising :

- (i) administering a test compound to a mammal having ovarian cancer or a cell isolated therefrom;
- (ii) comparing the level of gene expression of a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-3 in a treated cell or mammal with the level of gene expression of the polynucleotide in a control cell or mammal, wherein a test compound that

30

35

modulates the level of expression of the polynucleotide is a candidate for the treatment of ovarian cancer.

Typically, the control is a mammal with ovarian cancer or a cell therefrom that has not  
5 been treated with the test compound. Alternatively, the control is a normal cell or mammal.

The present invention also provides a method for treating a mammal having ovarian cancer comprising administering a compound identified the drug screening method  
10 *supra*.

In a further aspect, the present invention provides a pharmaceutical composition for use in treating a mammal having ovarian cancer, the composition comprising a compound identified the screening method *supra* for identifying a compound that modulates an  
15 ovarian cancer-associated polypeptide, or alternatively, using the drug screening method *supra*, and a physiologically acceptable carrier or diluent.

In a further aspect, the present invention provides an assay device, preferably for use in the diagnosis or prognosis of ovarian cancer, said device comprising a plurality of  
20 polynucleotides immobilized to a solid phase, wherein each of said polynucleotides consists of a gene as listed in any one of Tables 1-3. Preferably, the solid phase is a substantially planar chip.

In a related embodiment, the present invention provides an assay device, preferably for  
25 use in the diagnosis or prognosis of ovarian cancer, said device comprising a plurality of different antibodies immobilized to a solid phase, wherein each of said antibodies binds to a polypeptide listed in Tables 1-3. Preferably, the solid phase is a substantially planar chip.

30 Preferably, the assay device *supra* is used in a method of diagnosis or prognosis as described herein.

Alternatively, the assay device is used to identify modulatory compounds of the expression of one or more genes/proteins listed in any one of Tables 1-3.

In a further aspect, the present invention provides a non-human transgenic animal which is transgenic by virtue of comprising a gene set forth in any one of Tables 1-3 and, in particular, to the use of any such transgenic animal in the performance of a diagnostic or prognostic method of the invention as transgenic "knock-out" animals that have disrupted expression of a gene as set forth in any one of Tables 1-3.

In a further aspect, the present invention provides an isolated polynucleotide selected from the group consisting of:

- (a) polynucleotides comprising a nucleotide sequence as shown in Tables 1-3, or the complement thereof;
- (b) polynucleotides comprising a nucleotide sequence capable of selectively hybridizing to a nucleotide sequence as shown in Tables 1-3;
- (c) polynucleotides comprising a nucleotide sequence capable of selectively hybridizing to the complement of a nucleotide sequence as shown in Tables 1-3;
- and
- (d) polynucleotides comprising a polynucleotide sequence which is degenerate as a result of the genetic code to the polynucleotides defined in (a), (b) or (c).

Preferred polynucleotides comprise a polynucleotide sequence as shown in Tables 1-3 or a sequence having at least 80% homology thereto.

Preferably, the isolated polynucleotide is used for the diagnosis or prognosis of ovarian cancer, more preferably by a method as described herein. In a particularly preferred embodiment, the present invention provides for the use of a polynucleotide as set forth in any one of Tables 1-3 in the diagnosis or prognosis of ovarian cancer or for the preparation of a medicament for the treatment of ovarian cancer.

The present invention also provides a nucleic acid vector comprising a polynucleotide of the invention. In one embodiment, the polynucleotide is operably linked to a regulatory control sequence capable of directing expression of the polynucleotide in a host cell. In a particularly preferred embodiment, the present invention provides for the use of a vector comprising a polynucleotide as set forth in any one of Tables 1-3 in the diagnosis or prognosis of ovarian cancer or for the preparation of a medicament for the treatment of ovarian cancer.

The present invention further provides a host cell comprising a vector as described in the preceding paragraph. In a particularly preferred embodiment, the present invention provides for the use of a host cell comprising an introduced polynucleotide as set forth in any one of Tables 1-3 in the diagnosis or prognosis of ovarian cancer or for the preparation of a medicament for the treatment of ovarian cancer.

In a further aspect, the present invention provides an isolated polypeptide which is encoded by a gene set forth in any one of Tables 1-3. The present invention also provides an isolated polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-3. In a particularly preferred embodiment, the present invention provides for the use of an isolated polypeptide as set forth in any one of Tables 1-3 in the diagnosis or prognosis of ovarian cancer or for the preparation of a medicament for the treatment of ovarian cancer.

In a further aspect the present invention provides an antibody that binds specifically a polypeptide listed in Tables 1-3. In a particularly preferred embodiment, the present invention provides for the use of an antibody that binds to an isolated polypeptide as set forth in any one of Tables 1-3 in the diagnosis or prognosis of ovarian cancer or for the preparation of a medicament for the treatment of ovarian cancer.

20

#### Brief description of the Drawings

Figure 1 is a photographic representation showing expression of genes as identified by immunohistochemical staining of fixed normal (i.e. non-cancerous or healthy) tissues (panel A) or ovarian cancer tissue (panel B). The inset in panel A shows inclusion cysts. The expression levels of the following genes listed in Table 1 or Table 3 were determined: Claudin-3 (SEQ ID NO: 15); EP-CAM (Accession No. NM\_002354); and SOX17 (SEQ ID NO: 17). Positive controls CA125, MUC-1 and E-Cadherin were also included for comparison.



Figure 2 is a graphical representation showing the correlation between expression of different genes in serous ovarian cancer (SOC), mucinous ovarian cancer (MOC), endometrioid ovarian cancer (EnOC) and clear cell ovarian cancer (CICA). Genes indicated on the x-axis in each case are as in the legend to Figure 1.

Figure 3 is a copy of a photographic representation showing immunohistochemical staining of ovary tissue from a normal healthy subject (normal ovary), a subject diagnosed with mucinous ovarian cancer (MOC) and a subject diagnosed with serous ovarian cancer (SOC), following staining with probes that are specific for L1-Cadherin (top row), meprin alpha (middle row) or galectin-4 (lower row). Magnification is indicated as 20-40X.

Figure 4a is a copy of a photographic representation showing immunohistochemical staining of samples from a normal healthy subject (normal) or primary serous ovarian tumor (SOC), following staining with probes that are specific for sFRP4 (top row), or SOCS3 (lower row). Magnification is indicated as 20X.

Figure 4b is a copy of a graphical representation showing a Kaplan-Meier survival curve correlating sFRP4 expression to patient survival over the medium term (i.e., from about 12 months to about 48 months) to long term (more than about 48 months), indicating that high expression of sFRP4 is associated with poor survival in patients (n=127) having SOC (p=0.0056).

#### Detailed description of the preferred embodiments

##### *Ovarian cancer-associated sequences:*

Ovarian cancer-associated sequences can include both nucleic acid (i.e., "ovarian cancer-associated genes") and protein (i.e., "ovarian cancer-associated proteins").

As used herein, the term "ovarian cancer-associated protein" shall be taken to mean any protein that has an expression pattern correlated to an ovarian cancer, the recurrence of an ovarian cancer or the survival of a subject suffering from ovarian cancer.

Similarly, the term "ovarian cancer-associated gene" shall be taken to mean any nucleic acid encoding an ovarian cancer-associated protein or nucleic acid having an expression profile that is correlated to an ovarian cancer, the recurrence of an ovarian cancer or the survival of a subject suffering from ovarian cancer.

5

As will be appreciated by those in the art and is more fully outlined below, ovarian cancer-associated genes are useful in a variety of applications, including diagnostic applications, which will detect naturally occurring nucleic acids, as well as screening applications; e.g., biochips comprising nucleic acid probes or PCR microtitre plates with selected probes to the ovarian cancer sequences are generated.

10

For identifying ovarian cancer-associated sequences, the ovarian cancer screen typically includes comparing genes identified in different tissues, e.g., normal and cancerous tissues, or tumour tissue samples from patients who have metastatic disease vs. non metastatic tissue. Other suitable tissue comparisons include comparing ovarian cancer samples with metastatic cancer samples from other cancers, such as lung, breast, gastrointestinal cancers, ovarian, etc. Samples of different stages of ovarian cancer, e.g., survivor tissue, drug resistant states, and tissue undergoing metastasis, are applied to biochips comprising nucleic acid probes. The samples are first microdissected, if applicable, and treated as is known in the art for the preparation of mRNA. Suitable blochips are commercially available, e.g. from Affymetrix. Gene expression profiles as described herein are generated and the data analyzed.

15

20

In one embodiment, the genes showing changes in expression as between normal and disease states are compared to genes expressed in other normal tissues, preferably normal ovarian, but also including, and not limited to lung, heart, brain, liver, breast, kidney, muscle, colon, small intestine, large intestine, spleen, bone and placenta. In a preferred embodiment, those genes identified during the ovarian cancer screen that are expressed in any significant amount in other tissues are removed from the profile, although in some embodiments, this is not necessary. That is, when screening for drugs, it is usually preferable that the target be disease specific, to minimise possible side effects.

30

In a preferred embodiment, ovarian cancer-associated sequences are those that are up-regulated in ovarian cancer; that is, the expression of these genes is modified (up-

35

regulated or down-regulated) in ovarian cancer tissue as compared to non-cancerous tissue (see Table 1).

"Up-regulation" as used herein means at least about a two-fold change, preferably at least about a three fold change, with at least about five-fold or higher being preferred. All Unigene cluster identification numbers and accession numbers herein are for the GenBank sequence database and the sequences of the accession numbers are hereby expressly incorporated by reference. Sequences are also available in other databases, e.g., European Molecular Biology Laboratory (EMBL) and DNA Database of Japan (DDBJ).

"Down-regulation" as used herein often means at least about a 1.5-fold change more preferably a two-fold change, preferably at least about a three fold change, with at least about five-fold or higher being most preferred.

Particularly preferred sequences are those referred to in Tables 1 or 3 that have a P value of less than 0.05, more preferably a P value of less than about 0.01.

Similarly, preferred sequences are those referred to in Table 2 as having an absolute ratio of expression between ovarian patients and normal patients of at least about  $\pm 5.0$ , more preferably at least about  $\pm 6.0$ , even more preferably at least about  $\pm 7.0$  or at least about  $\pm 8.0$  or at least about  $\pm 9.0$  or at least about  $\pm 10.0$ .

*Detection of ovarian cancer sequences for diagnostic/prognostic applications*

In one aspect, the RNA expression levels of genes are determined for different cellular states in the ovarian cancer phenotype. Expression levels of genes in normal tissue (i.e., not undergoing ovarian cancer) and in ovarian cancer tissue (and in some cases, for varying severities of ovarian cancer that relate to prognosis, as outlined below) are evaluated to provide expression profiles. An expression profile of a particular cell state or point of development is essentially a "fingerprint" of the state. While two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is reflective of the state of the cell. By comparing expression profiles of cells in different states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained. Then, diagnosis are performed or confirmed to

determine whether a tissue sample has the gene expression profile of normal or cancerous tissue. This will provide for molecular diagnosis of related conditions.

"Differential expression," or grammatical equivalents as used herein, refers to qualitative or quantitative differences in the temporal and/or cellular gene expression patterns within and among cells and tissue. Thus, a differentially expressed gene can qualitatively have its expression altered, including an activation or inactivation, in, e.g., normal versus ovarian cancer tissue. Genes are turned on or turned off in a particular state, relative to another state thus permitting comparison of two or more states. A qualitatively regulated gene will exhibit an expression pattern within a state or cell type which is detectable by standard techniques. Some genes will be expressed in one state or cell type, but not in both. Alternatively, the difference in expression are quantitative, e.g., in that expression is increased or decreased; i.e., gene expression is either upregulated, resulting in an increased amount of transcript, or downregulated, resulting in a decreased amount of transcript. The degree to which expression differs need only be large enough to quantify via standard characterization techniques as outlined below, such as by use of Affymetrix GeneChip<sup>TM</sup> expression arrays, Lockhart, *Nature Biotechnology* 14:1675-1680 (1996), hereby expressly incorporated by reference. Other techniques include, but are not limited to, quantitative reverse transcriptase PCR, northern analysis and RNase protection. As outlined above, preferably the change in expression (i.e., upregulation or downregulation) is at least about 50%, more preferably at least about 100%, more preferably at least about 150%, more preferably at least about 200%, with from 300 to at least 1000% being especially preferred.

Evaluation are at the gene transcript, or the protein level. The amount of gene expression are monitored using nucleic acid probes to the DNA or RNA equivalent of the gene transcript, and the quantification of gene expression levels, or, alternatively, the final gene product itself (protein) are monitored, e.g., with antibodies to the ovarian cancer-associated protein and standard immunoassays (ELISAs, etc.) or other techniques, including mass spectroscopy assays, 2D gel electrophoresis assays, etc. Proteins corresponding to ovarian cancer genes, i.e., those identified as being important in a ovarian cancer phenotype, are evaluated in a ovarian cancer diagnostic test.

In a preferred embodiment, gene expression monitoring is performed on a plurality of genes. Multiple protein expression monitoring are performed as well. Similarly, these assays are performed on an individual basis as well.

In this embodiment, the ovarian cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of ovarian cancer sequences in a particular cell. The assays are further described below in the example. PCR techniques  
5 are used to provide greater sensitivity.

In a preferred embodiment nucleic acids encoding the ovarian cancer-associated protein are detected. Although DNA or RNA encoding the ovarian cancer-associated protein are detected, of particular interest are methods wherein an mRNA encoding a ovarian  
10 cancer-associated protein is detected. Probes to detect mRNA are a nucleotide/deoxynucleotide probe that is complementary to and hybridizes with the mRNA and includes, but is not limited to, oligonucleotides, cDNA or RNA. Probes also should contain a detectable label, as defined herein. In one method the mRNA is detected after immobilizing the nucleic acid to be examined on a solid support such as  
15 nylon membranes and hybridizing the probe with the sample. Following washing to remove the non-specifically bound probe, the label is detected. In another method detection of the mRNA is performed in situ. In this method permeabilized cells or tissue samples are contacted with a detectably labeled nucleic acid probe for sufficient time to allow the probe to hybridize with the target mRNA. Following washing to remove the  
20 non-specifically bound probe, the label is detected. For example a digoxigenin labeled riboprobe (RNA probe) that is complementary to the mRNA encoding a ovarian cancer-associated protein is detected by binding the digoxigenin with an anti-digoxigenin secondary antibody and developed with nitro blue tetrazolium and 5-bromo-4-chloro-3-indoyl phosphate.

In a preferred embodiment, various proteins from the three classes of proteins as described herein (secreted, transmembrane or intracellular proteins) are used in diagnostic assays. The ovarian cancer-associated proteins, antibodies, nucleic acids, modified proteins and cells containing ovarian cancer sequences are used in diagnostic  
25 assays. This are performed on an individual gene or corresponding polypeptide level. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes and/or corresponding polypeptides.

As described and defined herein, ovarian cancer-associated proteins, including  
35 intracellular, transmembrane or secreted proteins, find use as markers of ovarian cancer.

Detection of these proteins in putative ovarian cancer tissue allows for detection or diagnosis of ovarian cancer. In one embodiment, antibodies are used to detect ovarian cancer-associated proteins. A preferred method separates proteins from a sample by electrophoresis on a gel (typically a denaturing and reducing protein gel, but are another  
5 type of gel, including isoelectric focusing gels and the like). Following separation of proteins, the ovarian cancer-associated protein is detected, e.g., by immunoblotting with antibodies raised against the ovarian cancer-associated protein. Methods of immunoblotting are well known to those of ordinary skill in the art.

10 In another preferred method, antibodies to the ovarian cancer-associated protein find use in *in situ* imaging techniques, e.g., in histology (e.g., *Methods in Cell Biology: Antibodies in Cell Biology*, volume 37 (Asai, ed. 1993)). In this method cells are contacted with from one to many antibodies to the ovarian cancer-associated protein(s). Following washing to remove non-specific antibody binding, the presence of the antibody or antibodies is  
15 detected. In one embodiment the antibody is detected by incubating with a secondary antibody that contains a detectable label. In another method the primary antibody to the ovarian cancer-associated proteins) contains a detectable label, e.g. an enzyme marker that can act on a substrate. In another preferred embodiment each one of multiple primary antibodies contains a distinct and detectable label. This method finds particular  
20 use in simultaneous screening for a plurality of ovarian cancer-associated proteins. As will be appreciated by one of ordinary skill in the art, many other histological imaging techniques are also provided by the invention.

In a preferred embodiment the label is detected in a fluorometer which has the ability to  
25 detect and distinguish emissions of different wavelengths. In addition, a fluorescence activated cell sorter (FACS) are used in the method. In another preferred embodiment, antibodies find use in diagnosing ovarian cancer from blood, serum, plasma, stool, and other samples. Such samples, therefore, are useful as samples to be probed or tested for the presence of ovarian cancer-associated proteins. Antibodies are used to detect a  
30 ovarian cancer-associated protein by previously described immunoassay techniques including ELISA, immunoblotting (western blotting), immunoprecipitation, BIACORE technology and the like. Conversely, the presence of antibodies may indicate an immune response against an endogenous ovarian cancer-associated protein.

35 In a preferred embodiment, *in situ* hybridization of labeled ovarian cancer nucleic acid probes to tissue arrays is done. For example, arrays of tissue samples, including ovarian

cancer tissue and/or normal tissue, are made. In situ hybridization (see, e.g., Ausubel, supra) is then performed. When comparing the fingerprints between an individual and a standard, the skilled artisan can make a diagnosis, a prognosis, or a prediction based on the findings. It is further understood that the genes which indicate the diagnosis may differ from those which indicate the prognosis and molecular profiling of the condition of the cells may lead to distinctions between responsive or refractory conditions or are predictive of outcomes.

In a preferred embodiment, the ovarian cancer-associated proteins, antibodies, nucleic acids, modified proteins and cells containing ovarian cancer sequences are used in prognosis assays. As above, gene expression profiles are generated that correlate to ovarian cancer, in terms of long term prognosis. Again, this are done on either a protein or gene level, with the use of genes being preferred. As above, ovarian cancer probes are attached to biochips for the detection and quantification of ovarian cancer sequences in a tissue or patient. The assays proceed as outlined above for diagnosis. PCR method may provide more sensitive and accurate quantification.

*Characteristics of ovarian cancer-associated proteins and genes encoding same*

Ovarian cancer-associated proteins of the present invention are classified as secreted proteins, transmembrane proteins or intracellular proteins. In one embodiment, the ovarian cancer-associated protein is an intracellular protein. Intracellular proteins are found in the cytoplasm and/or in the nucleus. Intracellular proteins are involved in all aspects of cellular function and replication (including, e.g., signaling pathways); aberrant expression of such proteins often results in unregulated or dysregulated cellular processes (see, e.g., Molecular Biology of the Cell (Alberts, ed., 3rd ed., 1994). For example, many intracellular proteins have enzymatic activity such as protein kinase activity, protein phosphatase activity, protease activity, nucleotide cyclase activity, polymerase activity and the like. Intracellular proteins also serve as docking proteins that are involved in organizing complexes of proteins, or targeting proteins to various subcellular localizations, and are involved in maintaining the structural integrity of organelles.

An increasingly appreciated concept in characterising proteins is the presence in the proteins of one or more motifs for which defined functions have been attributed. In addition to the highly conserved sequences found in the enzymatic domain of proteins, highly conserved sequences have been identified in proteins that are involved in

protein-protein interaction. For example, Src-homology-2 (SH2) domains bind tyrosine-phosphorylated targets in a sequence dependent manner. PTB domains, which are distinct from SH2 domains, also bind tyrosine phosphorylated targets. SH3 domains bind to proline-rich targets. In addition, PH domains, tetratricopeptide repeats and WD domains to name only a few, have been shown to mediate protein-protein interactions. Some of these may also be involved in binding to phospholipids or other second messengers. As will be appreciated by one of ordinary skill in the art, these motifs are identified on the basis of primary sequence; thus, an analysis of the sequence of proteins may provide insight into both the enzymatic potential of the molecule and/or molecules with which the protein may associate. One useful database is Pfam (protein families), which is a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains. Versions are available via the internet from Washington University in St. Louis, the Sanger Center in England, and the Karolinska Institute in Sweden (see, e.g., Bateman et al., 2000, *Nuc. Acids Res.* 28: 263-266; Sonnhammer et al., 1997, *Proteins* 28: 405-420; Bateman et al., 1999, *Nuc. Acids Res.* 27:260-262; and Sonnhammer et al., 1998, *Nuc. Acids Res.* 26: 320-322.

In another embodiment, the ovarian cancer sequences are transmembrane proteins. Transmembrane proteins are molecules that span a phospholipid bilayer of a cell. They may have an intracellular domain, an extracellular domain, or both. The intracellular domains of such proteins may have a number of functions including those already described for intracellular proteins. For example, the intracellular domain may have enzymatic activity and/or may serve as a binding site for additional proteins. Frequently the intracellular domain of transmembrane proteins serves both roles. For example certain receptor tyrosine kinases have both protein kinase activity and SH2 domains. In addition, autophosphorylation of tyrosines on the receptor molecule itself, creates binding sites for additional SH2 domain containing proteins.

Transmembrane proteins may contain from one to many transmembrane domains. For example, receptor tyrosine kinases, certain cytokine receptors, receptor guanylyl cyclases and receptor serine/threonine protein kinases contain a single transmembrane domain. However, various other proteins including channels and adenylyl cyclases contain numerous transmembrane domains. Many important cell surface receptors such as G protein coupled receptors (GPCRs) are classified as "seven transmembrane domain" proteins, as they contain 7 membrane spanning regions. Characteristics of transmembrane domains include approximately 20 consecutive hydrophobic amino acids



- that are followed by charged amino acids. Therefore, upon analysis of the amino acid sequence of a particular protein, the localization and number of transmembrane domains within the protein are predicted (see, e.g. PSORT web site <http://psort.nibb.ac.jp/>). Important transmembrane protein receptors include, but are not limited to the insulin
- 5 receptor, insulin-like growth factor receptor, human growth hormone receptor, glucose transporters, transferrin receptor, epidermal growth factor receptor, low density lipoprotein receptor, epidermal growth factor receptor, leptin receptor, interleukin receptors, e.g. IL-1 receptor, IL-2 receptor,
- 10 The extracellular domains of transmembrane proteins are diverse; however, conserved motifs are found repeatedly among various extracellular domains. Conserved structure and/or functions have been ascribed to different extracellular motifs. Many extracellular domains are involved in binding to other molecules. In one aspect, extracellular domains are found on receptors. Factors that bind the receptor domain include circulating ligands,
- 15 which are peptides, proteins, or small molecules such as adenosine and the like. For example, growth factors such as EGF, FGF and PDGF are circulating growth factors that bind to their cognate receptors to initiate a variety of cellular responses. Other factors include cytokines, mitogenic factors, neurotrophic factors and the like. Extracellular domains also bind to cell-associated molecules. In this respect, they mediate cell-cell
- 20 interactions., Cell-associated ligands are tethered to the cell, e.g., via a glycosylphosphatidylinositol (GPI) anchor, or may themselves be transmembrane proteins. Extracellular domains also associate with the extracellular matrix and contribute to the maintenance of the cell structure.
- 25 Ovarian cancer-associated proteins that are transmembrane are particularly preferred in the present invention as they are readily accessible targets for immunotherapeutics, as are described herein. In addition, as outlined below, transmembrane proteins are also useful in imaging modalities. Antibodies are used to label such readily accessible proteins *in situ*. Alternatively, antibodies can also label intracellular proteins, in which
- 30 case samples are typically permeabilized to provide access to intracellular proteins.
- It will also be appreciated by those in the art that a transmembrane protein are made soluble by removing transmembrane sequences, e.g., through recombinant methods. Furthermore, transmembrane proteins that have been made soluble are made to be
- 35 secreted through recombinant means by adding an appropriate signal sequence.

In another embodiment, the ovarian cancer-associated proteins are secreted proteins; the secretion of which are either constitutive or regulated. These proteins have a signal peptide or signal sequence that targets the molecule to the secretory pathway. Secreted proteins are involved in numerous physiological events; by virtue of their circulating nature, they serve to transmit signals to various other cell types. The secreted protein may function in an autocrine manner (acting on the cell that secreted the factor), a paracrine manner (acting on cells in close proximity to the cell that secreted the factor) or an endocrine manner (acting on cells at a distance). Thus secreted molecules find use in modulating or altering numerous aspects of physiology. Ovarian cancer-associated proteins that are secreted proteins are particularly preferred in the present invention as they serve as good targets for diagnostic markers, e.g., for blood, plasma, serum, or stool tests.

#### *Mammalian subjects*

The present invention provides nucleic acid and protein sequences that are differentially expressed in ovarian cancer, herein termed "ovarian cancer sequences." As outlined below, ovarian cancer sequences include those that are up-regulated (i.e., expressed at a higher level) in ovarian cancer, as well as those that are down-regulated (i.e., expressed at a lower level). In a preferred embodiment, the ovarian cancer sequences are from humans; however, as will be appreciated by those in the art, ovarian cancer sequences from other organisms are useful in animal models of disease and drug evaluation; thus, other ovarian cancer sequences are provided, from vertebrates, including mammals, including rodents (rats, mice, hamsters, guinea pigs, etc.), primates, farm animals (including sheep, goats, pigs, cows, horses, etc.) and pets, e.g., (dogs, cats, etc.).

#### *Assay control samples*

It will be apparent from the preceding discussion that many of the diagnostic methods provided by the present invention involve a degree of quantification to determine, on the one hand, the over-expression or reduced-expression of a diagnostic/prognostic marker in tissue that is suspected of comprising a cancer cell. Such quantification can be readily provided by the inclusion of appropriate control samples in the assays described below, derived from healthy or normal individuals. Alternatively, if internal controls are not included in each assay conducted, the control may be derived from an established data set that has been generated from healthy or normal individuals.

In the present context, the term "healthy individual" shall be taken to mean an individual who is known not to suffer from ovarian cancer, such knowledge being derived from clinical data on the individual, including, but not limited to, a different cancer assay to that described herein. As the present invention is particularly useful for the early detection of ovarian cancer, it is preferred that the healthy individual is asymptomatic with respect to the early symptoms associated with ovarian cancer. Although early detection using well-known procedures is difficult, reduced urinary frequency, rectal pressure, and abdominal bloating and swelling, are associated with the disease in its early stages, and, as a consequence, healthy individuals should not have any of these clinical symptoms. Clearly, subjects suffering from later symptoms associated with ovarian cancer, such as, for example, metastases in the omentum, abdominal fluid, lymph nodes, lung, liver, brain, or bone, and subjects suffering from spinal cord compression, elevated calcium level, chronic pain, or pleural effusion, should also be avoided from the "healthy individual" data set.

The term "normal individual" shall be taken to mean an individual having a normal level of expression of a cancer-associate gene or cancer-associated protein in a particular sample derived from said individual. As will be known to those skilled in the art, data obtained from a sufficiently large sample of the population will normalize, allowing the generation of a data set for determining the average level of a particular parameter. Accordingly, the level of expression of a cancer-associate gene or cancer-associated protein can be determined for any population of individuals, and for any sample derived from said individual, for subsequent comparison to levels determined for a sample being assayed. Where such normalized data sets are relied upon, internal controls are preferably included in each assay conducted to control for variation.

In one embodiment, the present invention provides a method for detecting a cancer cell in a subject, said method comprising:

- (i) determining the level of mRNA encoding a cancer-associated protein expressed in a test sample from said subject; and
- (ii) comparing the level of mRNA determined at (i) to the level of mRNA encoding a cancer-associated protein expressed in a comparable sample from a healthy or normal individual,

wherein a level of mRNA at (i) that is modified in the test sample relative to the comparable sample from the normal or healthy individual is indicative of the presence of a cancer cell in said subject.

Alternatively, or in addition, the controll may comprise a cancer-associated sequence that is known to be expressed at a particular level in an ovarian cancer, eg., CA125, MUC-1 or E-Cadherin, amongst others.

5

*Biological samples*

Preferred biological samples in which the assays of the invention are performed include bodily fluids, ovarian tissue and cells, and those tissues known to comprise cancer cells arising from a metastasis of an ovarian cancer, such as, for example, in carcinomas of  
10 the lung, prostate, breast, colon, pancreas, placenta, or omentum, and in cells of brain anaplastic oligodendrogliomas.

Bodily fluids shall be taken to include whole blood, serum, peripheral blood mononuclear cells (PBMC), or buffy coat fraction.

15

In the present context, the term "cancer cell" includes any biological specimen or sample comprising a cancer cell irrespective of its degree of isolation or purity, such as, for example, tissues, organs, cell lines, bodily fluids, or histology specimens that comprise a cell in the early stages of transformation or having been transformed.

20

As the present invention is particularly useful for the early detection and prognosis of cancer of the medium to long term, the definition of "cancer cell" is not to be limited by the stage of a cancer in the subject from which said cancer cell is derived (ie. whether or not the patient is in remission or undergoing disease recurrence or whether or not the  
25 cancer is a primary tumor or the consequence of metastases). Nor is the term "cancer cell" to be limited by the stage of the cell cycle of said cancer cell.

Preferably, the sample comprises ovarian tissue, prostate tissue, kidney tissue, uterine tissue, placenta, a cervical specimen, omentum, rectal tissue, brain tissue, bone tissue,  
30 lung tissue, lymphatic tissue, urine, semen, blood, abdominal fluid, or serum, or a cell preparation or nucleic acid preparation derived therefrom. More preferably, the sample comprises serum or abdominal fluid, or a tissue selected from the group consisting of: ovary, lymph, lung, liver, brain, placenta, brain, omentum, and prostate. Even more preferably, the sample comprises serum or abdominal fluid, ovary (eg. OSE), or lymph  
35 node tissue. The sample can be prepared on a solid matrix for histological analyses, or alternatively, in a suitable solution such as, for example, an extraction buffer or

suspension buffer, and the present invention clearly extends to the testing of biological solutions thus prepared.

*Polynucleotide probes and amplification primers*

5 Polynucleotide probes are derived from or comprise the nucleic acid sequences whose nucleotide sequences are provided by reference to the public database accession numbers given in Tables 1-3 (referred to herein as the nucleotide sequences shown in Tables 1-3), and sequences homologues thereto as well as variants, derivatives and fragments thereof.

10

Whilst the probes may comprise double-stranded or single-stranded nucleic acid, single-stranded probes are preferred because they do not require melting prior to use in hybridizations. On the other hand, longer probes are also preferred because they can be used at higher hybridization stringency than shorter probes and may produce lower background hybridization than shorter probes.

15

So far as shorter probes are concerned, single-stranded, chemically-synthesized oligonucleotide probes are particularly preferred by the present invention. To reduce the noise associated with the use of such probes during hybridization, the nucleotide sequence of the probe is carefully selected to maximize the  $T_m$  at which hybridizations can be performed, reduce non-specific hybridization, and to reduce self-hybridization. Such considerations may be particularly important for applications involving high throughput screening using microarray technology. In general, this means that the nucleotide sequence of an oligonucleotide probe is selected such that it is unique to the target RNA or protein-encoding sequence, has a low propensity to form secondary structure, low self-complementary, and is not highly A/T-rich.

20

25

The only requirement for the probes is that they cross-hybridize to nucleic acid encoding the target diagnostic protein or the complementary nucleotide sequence thereto and are sufficiently unique in sequence to generate high signal:noise ratios under specified hybridization conditions. As will be known to those skilled in the art, long nucleic acid probes are preferred because they tend to generate higher signal:noise ratios than shorter probes and/or the duplexes formed between longer molecules have higher melting temperatures (i.e.  $T_m$  values) than duplexes involving short probes. Accordingly, full-length DNA or RNA probes are contemplated by the present invention, as are specific probes comprising the sequence of the 3'-untranslated region or complementary thereto.

30

35

In a particularly preferred embodiment, the nucleotide sequence of an oligonucleotide probe has no detectable nucleotide sequence identity to a nucleotide sequence in a BLAST search (Altschul *et al.*, *J. Mol. Biol.* 215, 403-410, 1990) or other database search, other than a sequence selected from the group consisting of: (a) a sequence encoding a polypeptide listed in any one of Tables 1-3; (b) the 5'-untranslated region of a sequence encoding a polypeptide listed in any one of Tables 1-3; (c) a 3'-untranslated region of a sequence encoding a polypeptide listed in any one of Tables 1-3; and (d) an exon region of a sequence encoding a polypeptide listed in any one of Tables 1-3.

10

Additionally, the self-complementarity of a nucleotide sequence can be determined by aligning the sequence with its reverse complement, wherein detectable regions of identity are indicative of potential self-complementarity. As will be known to those skilled in the art, such sequences may not necessarily form secondary structures during hybridization reaction, and, as a consequence, successfully identify a target nucleotide sequence. It is also known to those skilled in the art that, even where a sequence does form secondary structures during hybridization reactions, reaction conditions can be modified to reduce the adverse consequences of such structure formation. Accordingly, a potential for self-complementarity should not necessarily exclude a particular candidate oligonucleotide from selection. In cases where it is difficult to determine nucleotide sequences having no potential self-complementarity, the uniqueness of the sequence should outweigh a consideration of its potential for secondary structure formation.

15

20

Recommended pre-requisites for selecting oligonucleotide probes, particularly with respect to probes suitable for microarray technology, are described in detail by Lockhart *et al.*, "Expression monitoring by hybridization to high-density oligonucleotide arrays", *Nature Biotech.* 14, 1675-1680, 1996.

25

The nucleic acid probe may comprise a nucleotide sequence that is within the coding strand of a gene listed in any one of Tables 1-3. Such "sense" probes are useful for detecting RNA by amplification procedures, such as, for example, polymerase chain reaction (PCR), and more preferably, quantitative PCR or reverse transcription polymerase chain reaction (RT-PCR). Alternatively, "sense" probes may be expressed to produce polypeptides or immunologically active derivatives thereof that are useful for detecting the expressed protein in samples.

30

35

The nucleotide sequences referred to in Tables 1-3 and homologues thereof, typically encode polypeptides. It will be understood by a skilled person that numerous different polynucleotides can encode the same polypeptide as a result of the degeneracy of the genetic code. In addition, it is to be understood that skilled persons may, using routine techniques, make nucleotide substitutions that do not affect the polypeptide sequence encoded by the polynucleotides of the invention to reflect the codon usage of any particular host organism in which the polypeptides of the invention are to be expressed.

- 5
- 10 Polynucleotides may comprise DNA or RNA. They are single-stranded or double-stranded. They may also be polynucleotides which include within them synthetic or modified nucleotides. A number of different types of modification to oligonucleotides are known in the art. These include methylphosphonate and phosphorothioate backbones, addition of acridine or polylysine chains at the 3' and/or 5' ends of the molecule. For the
- 15 purposes of the present invention, it is to be understood that the polynucleotides described herein are modified by any method available in the art. Such modifications are carried out in order to enhance the *in vivo* activity or life span of the diagnostic/prognostic polynucleotides.
- 20 The terms "variant" or "derivative" in relation to the nucleotide sequences of the present invention include any substitution of, variation of, modification of, replacement of, deletion of or addition of one (or more) nucleic acid from or to the sequence provided that the resultant nucleotide sequence codes for a polypeptide having biological activity, preferably having substantially the same activity as the polypeptide sequences presented in the sequence
- 25 listings.

With respect to sequence homology, preferably there is at least 75%, more preferably at least 85%, more preferably at least 90% homology to a sequence shown in Tables 1-3 herein over a region of at least 20, preferably at least 25 or 30, for instance at least 40, 60,

30 100, 500, 1000 or more contiguous nucleotides. More preferably there is at least 95%, more preferably at least 98%, homology. In one embodiment, homologues are naturally occurring sequences, such as orthologues, tissue-specific isoforms and allelic variants.

Homology comparisons are conducted by eye, or more usually, with the aid of readily

35 available sequence comparison programs. These commercially available computer programs can calculate % homology between two or more sequences.

Percentage (%) homology are calculated over contiguous sequences, i.e. one sequence is aligned with the other sequence and each nucleotide in one sequence directly compared with the corresponding nucleotide in the other sequence, one base at a time. This is called an "ungapped" alignment. Typically, such ungapped alignments are performed only over a relatively short number of bases (for example less than 50 contiguous nucleotides).

Although this is a very simple and consistent method, it fails to take into consideration that, for example, in an otherwise identical pair of sequences, one insertion or deletion will cause the following nucleotides to be put out of alignment, thus potentially resulting in a large reduction in % homology when a global alignment is performed. Consequently, most sequence comparison methods are designed to produce optimal alignments that take into consideration possible insertions and deletions without penalising unduly the overall homology score. This is achieved by inserting "gaps" in the sequence alignment to try to maximise local homology.

However, these more complex methods assign "gap penalties" to each gap that occurs in the alignment so that, for the same number of identical amino acids, a sequence alignment with as few gaps as possible - reflecting higher relatedness between the two compared sequences - will achieve a higher score than one with many gaps. "Affine gap costs" are typically used that charge a relatively high cost for the existence of a gap and a smaller penalty for each subsequent residue in the gap. This is the most commonly used gap scoring system. High gap penalties will of course produce optimised alignments with fewer gaps. Most alignment programs allow the gap penalties to be modified. However, it is preferred to use the default values when using such software for sequence comparisons.

In determining whether or not two amino acid sequences fall within the stated defined percentage identity limits, those skilled in the art will be aware that it is necessary to conduct a side-by-side comparison of amino acid sequences. In such comparisons or alignments, differences will arise in the positioning of non-identical amino acid residues depending upon the algorithm used to perform the alignment. In the present context, references to percentage identities and similarities between two or more amino acid sequences shall be taken to refer to the number of identical and similar residues respectively, between said sequences as determined using any standard algorithm known to those skilled in the art. In particular, amino acid identities and similarities are calculated using the GAP program of the Computer Genetics Group, Inc., University



Research Park, Madison, Wisconsin, United States of America (Devereaux *et al*, *Nucl. Acids Res.* 12, 387-395, 1984), which utilizes the algorithm of Needleman and Wunsch *J. Mol. Biol.* 48, 443-453, 1970, or alternatively, the CLUSTAL W algorithm of Thompson *et al.*, *Nucl. Acids Res.* 22, 4673-4680, 1994, for multiple alignments, to maximize the  
5 number of identical/similar amino acids and to minimize the number and/or length of sequence gaps in the alignment.

A suitable computer program for carrying out such an alignment is the GCG Wisconsin Bestfit package (University of Wisconsin, U.S.A.; Devereux *et al.*, 1984, *Nucleic Acids*  
10 *Research* 12:387). The default scoring matrix has a match value of 10 for each identical nucleotide and -9 for each mismatch. The default gap creation penalty is -50 and the default gap extension penalty is -3 for each nucleotide.

Examples of other software than can perform sequence comparisons include, but are not  
15 limited to, the BLAST package (see Ausubel *et al.*, 1999 *ibid* – Chapter 18), FASTA (Atschul *et al.*, 1990, *J. Mol. Biol.*, 403-410) and the GENEWORKS suite of comparison tools. Both BLAST and FASTA are available for offline and online searching (see Ausubel *et al.*, 1999 *ibid*, pages 7-58 to 7-60). However it is preferred to use the GCG Bestfit program.

20 Once the software has produced an optimal alignment, it is possible to calculate % homology, preferably % sequence identity. The software typically does this as part of the sequence comparison and generates a numerical result.

25 A preferred sequence comparison program is the GCG Wisconsin Bestfit program described above.

The present invention also encompasses the use of nucleotide sequences that are capable of hybridizing selectively to the sequences presented herein, or any variant, fragment or  
30 derivative thereof, or to the complement of any of the above. Nucleotide sequences are preferably at least 15 nucleotides in length, more preferably at least 20, 30, 40 or 50 nucleotides in length.

The term "hybridization" as used herein shall include "the process by which a strand of  
35 nucleic acid joins with a complementary strand through base pairing" as well as the process of amplification as carried out in polymerase chain reaction technologies.

Polynucleotides capable of selectively hybridizing to the nucleotide sequences presented herein; or to their complement, will be generally at least 70%, preferably at least 80 or 90% and more preferably at least 95% or 98% homologous to the corresponding nucleotide sequences referred to in Tables 1-3 over a region of at least 20, preferably at least 25 or 30, for instance at least 40, 60, 100, 500, 1000 or more contiguous nucleotides.

The term "selectively hybridizable" means that the polynucleotide used as a probe is used under conditions where a target polynucleotide is found to hybridize to the probe at a level significantly above background. The background hybridization may occur because of other polynucleotides present, for example, in the cDNA or genomic DNA library being screening. In this event, background implies a level of signal generated by interaction between the probe and a non-specific DNA member of the library which is less than 10 fold, preferably less than 100 fold as intense as the specific interaction observed with the target DNA. The intensity of interaction are measured, for example, by radiolabelling the probe, e.g. with  $^{32}\text{P}$ .

Hybridization conditions are based on the melting temperature ( $T_m$ ) of the nucleic acid binding complex, as taught in Berger and Kimmel (1987, Guide to Molecular Cloning Techniques, Methods in Enzymology, Vol 152, Academic Press, San Diego CA), and confer a defined "stringency" as explained below.

For the purposes of defining the level of stringency, a high stringency hybridization is achieved using a hybridization buffer and/or a wash solution comprising the following:

- (i) a salt concentration that is equivalent to 0.1xSSC-0.2xSSC buffer or lower salt concentration;
- (ii) a detergent concentration equivalent to 0.1% (w/v) SDS or higher; and
- (iii) an incubation temperature of 55°C or higher.

Conditions for specifically hybridizing nucleic acid, and conditions for washing to remove non-specific hybridizing nucleic acid, are well understood by those skilled in the art. For the purposes of further clarification only, reference to the parameters affecting hybridization between nucleic acid molecules is found in Ausubel *et al.* (Current Protocols in Molecular Biology, Wiley Interscience, ISBN 047150338, 1992), which is herein incorporated by reference.

Maximum stringency typically occurs at about  $T_m - 5^\circ\text{C}$  ( $5^\circ\text{C}$  below the  $T_m$  of the probe); high stringency at about  $5^\circ\text{C}$  to  $10^\circ\text{C}$  below  $T_m$ ; intermediate stringency at about  $10^\circ\text{C}$  to  $20^\circ\text{C}$  below  $T_m$ ; and low stringency at about  $20^\circ\text{C}$  to  $25^\circ\text{C}$  below  $T_m$ . As will be understood by those of skill in the art, a maximum stringency hybridization are used to  
5 identify or detect identical polynucleotide sequences while an intermediate (or low) stringency hybridization are used to identify or detect similar or related polynucleotide sequences.

In a preferred aspect, the present invention covers nucleotide sequences that can hybridize  
10 to the nucleotide sequence of the present invention under stringent conditions (e.g.  $65^\circ\text{C}$  and  $0.1\times\text{SSC}$  { $1\times\text{SSC} = 0.15\text{ M NaCl}$ ,  $0.015\text{ M Na}_3\text{ Citrate pH } 7.0$ }).

Where the diagnostic/prognostic polynucleotide is double-stranded, both strands of the duplex, either individually or in combination, are encompassed by the present invention.  
15 Where the polynucleotide is single-stranded, it is to be understood that the complementary sequence of that polynucleotide is also included within the scope of the present invention.

Polynucleotides which are not 100% homologous to the sequences of the present invention but are useful in performing the diagnostic and/or prognostic assays of the invention by  
20 virtue of their ability to selectively hybridize to the target gene transcript, or to encode an immunologically cross-reactive protein to the target protein, are obtained in a number of ways, such as, for example by probing DNA libraries made from a range of individuals, for example individuals from different populations. In particular, given that that changes in the expression of diagnostic/prognostic cancer-associated genes correlate with ovarian cancer,  
25 characterisation of variant sequences in individuals suffering from ovarian cancer is used to identify variations in the sequences of ovarian-cancer associated genes (and proteins) that are predictive of and/or causative of ovarian cancer.

Accordingly the present invention provides methods of identifying sequence variants that  
30 are associated with ovarian cancer which methods comprise determining all or part of the nucleotide sequence of a gene referred to in Tables 1-3, derived from an individual suffering from ovarian cancer and comparing the sequence to that of the corresponding wild-type gene.

35 In addition, other viral/bacterial, or cellular homologues particularly cellular homologues found in mammalian cells (e.g. rat, mouse, bovine and primate cells), are obtained and such

homologues and fragments thereof in general will be capable of selectively hybridizing to the sequences shown in the sequence listing herein. Such sequences are obtained by probing cDNA libraries made from or genomic DNA libraries from other animal species, and probing such libraries with probes comprising all or part of the sequences referred to in  
5 Tables 1-3 under conditions of medium to high stringency. Similar considerations apply to obtaining species homologues and allelic variants of the nucleotide sequences referred to in Tables 1-3.

10 Variants and strain/species homologues may also be obtained using degenerate PCR which will use primers designed to target sequences within the variants and homologues encoding conserved amino acid sequences within the sequences of the present invention. Conserved sequences are predicted, for example, by aligning the amino acid sequences from several variants/homologues. Sequence alignments are performed using computer software known in the art. For example the GCG Wisconsin PileUp program is widely used.

15 The primers used in degenerate PCR will contain one or more degenerate positions and will be used at stringency conditions lower than those used for cloning sequences with single sequence primers against known sequences.

20 Alternatively, such polynucleotides are obtained by site-directed mutagenesis of characterised sequences, such as the sequences referred to in Tables 1-3. This are useful where for example silent codon changes are required to sequences to optimise codon preferences for a particular host cell in which the polynucleotide sequences are being expressed. Other sequence changes are desired in order to introduce restriction enzyme  
25 recognition sites, or to alter the property or function of the polypeptides encoded by the polynucleotides..

Polynucleotides comprising a diagnostic/prognostic cancer-associated gene are used to produce a primer by standard derivatization means, e.g. a PCR primer, a primer for an  
30 alternative amplification reaction, a probe e.g. labelled with a detectable label by conventional means using radioactive or non-radioactive labels, or the polynucleotides are cloned into vectors. Such primers, probes and other fragments will be at least 15, preferably at least 20, for example at least 25, 30 or 40 nucleotides in length. Preferred fragments are less than 5000, 2000, 1000, 500 or 200 nucleotides in length.

35

Polynucleotides such as a DNA polynucleotides and probes according to the invention are produced by recombinant or synthetic means, including cloning by standard techniques.

5 In general, primers will be produced by synthetic means, involving a step wise manufacture of the desired nucleic acid sequence one nucleotide at a time. Techniques for accomplishing this using automated techniques are readily available in the art.

10 Longer polynucleotides will generally be produced using recombinant means, for example using PCR (polymerase chain reaction) cloning techniques. This will involve making a pair of primers (e.g. of about 15 to 30 nucleotides) flanking a region of the sequence which it is desired to clone, bringing the primers into contact with mRNA or cDNA obtained from an animal or human cell, performing a polymerase chain reaction under conditions which bring about amplification of the desired region, isolating the amplified fragment (e.g. by purifying the reaction mixture on an agarose gel) and recovering the amplified DNA. The primers are  
15 designed to contain suitable restriction enzyme recognition sites so that the amplified DNA are cloned into a suitable cloning vector

Polynucleotide probes or primers preferably carry a detectable label. Suitable labels include radioisotopes such as  $^{32}\text{P}$  or  $^{35}\text{S}$ , enzyme labels, or other protein labels such as  
20 biotin. Such labels are added to polynucleotides or primers and are detected using by techniques known in the art.

Polynucleotide probes or primers labeled or unlabeled are also used by a person skilled in the art in nucleic acid-based tests for detecting or sequencing diagnostic/prognostic  
25 cancer-associated gene.

Such tests for detecting generally comprise bringing a biological sample containing DNA or RNA into contact with a probe comprising a polynucleotide probe or primer under at least low stringency hybridization conditions and detecting any duplex formed between  
30 the probe/primer and nucleic acid in the sample. Such detection are achieved using techniques such as PCR or by immobilising the probe on a solid support, removing nucleic acid in the sample which is not hybridized to the probe, and then detecting nucleic acid which has hybridized to the probe. Alternatively, the sample nucleic acid are immobilised on a solid support, and the amount of probe bound to such a support are  
35 detected. Suitable assay methods of this and other formats are found in for example W089/03891 and W090/13667.

Tests for sequencing nucleotides include bringing a biological sample containing target DNA or RNA into contact with a probe comprising a polynucleotide probe or primer under at least low stringency hybridization conditions and determining the sequence by, for example the Sanger dideoxy chain termination method (see Sambrook et al.).

Such a method generally comprises elongating, in the presence of suitable reagents, the primer by synthesis of a strand complementary to the target DNA or RNA and selectively terminating the elongation reaction at one or more of an A, C, G or T/U residue; allowing strand elongation and termination reaction to occur; separating out according to size the elongated products to determine the sequence of the nucleotides at which selective termination has occurred. Suitable reagents include a DNA polymerase enzyme, the deoxynucleotides dATP, dCTP, dGTP and dTTP, a buffer and ATP. Dideoxynucleotides are used for selective termination.

Tests for detecting or sequencing nucleotides in a biological sample are used as part of the methods of the invention for detecting ovarian cancer-associated transcripts and monitoring the efficacy of treatment of patients suffering from ovarian cancer as described in more detail herein.

The probes/primers may conveniently be packaged in the form of a test kit in a suitable container. In such kits the probe are bound to a solid support where the assay format for which the kit is designed requires such binding. The kit may also contain suitable reagents for treating the sample to be probed, hybridizing the probe to nucleic acid in the sample, control reagents, instructions, and the like.

Preferably, a kit of the invention comprises primers/probes suitable for selectively detecting a plurality of sequences, more preferably for selectively detecting a plurality of sequences that are listed in one or more of Tables 1-3 as having a P value of less than 0.05, more preferably a P value of less than 0.01. Similarly, a kit of the invention preferably comprises primers suitable for selectively detecting a plurality of sequences referred to in Table 1 or 2 or 3.

#### *Nucleic acid-based assay formats*

As discussed in detail below, the status of expression of a cancer-associated gene in patient samples may be analyzed by a variety protocols that are well known in the art

including *in situ* hybridization, northern blotting techniques, RT-PCR analysis (such as, for example, performed on laser capture microdissected samples), and microarray technology, such as, for example, using tissue microarrays probed with nucleic acid probes, or nucleic acid microarrays (ie. RNA microarrays or amplified DNA microarrays) microarrays probed with nucleic acid probes. All such assay formats are encompassed by the present invention.

For high throughput screening of large numbers of samples, such as, for example, public health screening of subjects, particularly human subjects, having a higher risk of developing cancer, microarray technology is a preferred assay format.

In accordance with such high throughput formats, techniques for producing immobilised arrays of DNA molecules have been described in the art. Generally, most prior art methods describe how to synthesise single-stranded nucleic acid molecule arrays, using for example masking techniques to build up various permutations of sequences at the various discrete positions on the solid substrate. U.S. Patent No. 5,837,832, the contents of which are incorporated herein by reference, describes an improved method for producing DNA arrays immobilised to silicon substrates based on very large scale integration technology. In particular, U.S. Patent No. 5,837,832 describes a strategy called "tiling" to synthesize specific sets of probes at spatially-defined locations on a substrate which are used to produced the immobilised DNA arrays. U.S. Patent No. 5,837,832 also provides references for earlier techniques that may also be used.

Thus DNA are synthesised *in situ* on the surface of the substrate. However, DNA may also be printed directly onto the substrate using for example robotic devices equipped with either pins or piezo electric devices.

The plurality of polynucleotide sequences are typically immobilised onto or in discrete regions of a solid substrate. The substrate are porous to allow immobilisation within the substrate or substantially non-porous, in which case the library sequences are typically immobilised on the surface of the substrate. The solid substrate are made of any material to which polypeptides can bind, either directly or indirectly. Examples of suitable solid substrates include flat glass, silicon wafers, mica, ceramics and organic polymers such as plastics, including polystyrene and polymethacrylate. It may also be possible to use semi-permeable membranes such as nitrocellulose or nylon membranes, which are widely available. The semi-permeable membranes are mounted on a more robust solid

surface such as glass. The surfaces may optionally be coated with a layer of metal, such as gold, platinum or other transition metal. A particular example of a suitable solid substrate is the commercially available BIAcore™ chip (Pharmacia Biosensors).

- 5 Preferably, the solid substrate is generally a material having a rigid or semi-rigid surface. In preferred embodiments, at least one surface of the substrate will be substantially flat, although in some embodiments it is desirable to physically separate synthesis regions for different polymers with, for example, raised regions or etched trenches. It is also preferred that the solid substrate is suitable for the high density application of DNA  
10 sequences in discrete areas of typically from 50 to 100  $\mu\text{m}$ , giving a density of 10000 to 40000  $\text{cm}^{-2}$ .

The solid substrate is conveniently divided up into sections. This is achieved by techniques such as photoetching, or by the application of hydrophobic inks, for example  
15 teflon-based inks (Cel-line, USA).

Discrete positions, in which each different member of the array is located may have any convenient shape, e.g., circular, rectangular, elliptical, wedge-shaped, etc.

- 20 Attachment of the polynucleotide sequences to the substrate are by covalent or non-covalent means. The plurality of polynucleotide sequences are attached to the substrate via a layer of molecules to which the sequences bind. For example, the sequences are labelled with biotin and the substrate coated with avidin and/or streptavidin. A convenient feature of using biotinylated sequences is that the efficiency of coupling to the  
25 solid substrate are determined easily. Since the library sequences may bind only poorly to some solid substrates, it is often necessary to provide a chemical interface between the solid substrate (such as in the case of glass) and the sequences. Examples of suitable chemical interfaces include hexaethylene glycol. Another example is the use of polylysine coated glass, the polylysine then being chemically modified using standard  
30 procedures to introduce an affinity ligand. Other methods for attaching molecules to the surfaces of solid substrate by the use of coupling agents are known in the art, see for example WO98/49557.

The complete DNA array is typically read at the same time by charged coupled device  
35 (CCD) camera or confocal imaging system. Alternatively, the DNA array are placed for detection in a suitable apparatus that can move in an x-y direction, such as a plate



reader. In this way, the change in characteristics for each discrete position are measured automatically by computer controlled movement of the array to place each discrete element in turn in line with the detection means.

- 5 The detection means are capable of interrogating each position in the library array optically or electrically. Examples of suitable detection means include CCD cameras or confocal imaging systems.

- 10 In a preferred embodiment, the level of expression of the cancer-associated gene in the test sample is determined by hybridizing a probe/primer to RNA in the test sample under at least low stringency hybridization conditions and detecting the hybridization using a detection means.

- 15 Similarly, the level of mRNA in the comparable sample from the healthy or normal individual is preferably determined by hybridizing a probe/primer to RNA in said comparable sample under at least low stringency hybridization conditions and detecting the hybridization using a detection means.

- 20 For the purposes of defining the level of stringency to be used in these diagnostic assays, a low stringency is defined herein as being a hybridization and/or a wash carried out in 6xSSC buffer, 0.1% (w/v) SDS at 28°C, or equivalent conditions. A moderate stringency is defined herein as being a hybridization and/or washing carried out in 2xSSC buffer, 0.1% (w/v) SDS at a temperature in the range 45°C to 65°C, or equivalent conditions. A high stringency is defined herein as being a hybridization and/or wash  
25 carried out in 0.1xSSC buffer, 0.1% (w/v) SDS, or lower salt concentration, and at a temperature of at least 65°C, or equivalent conditions. Reference herein to a particular level of stringency encompasses equivalent conditions using wash/hybridization solutions other than SSC known to those skilled in the art.

- 30 Generally, the stringency is increased by reducing the concentration of SSC buffer, and/or increasing the concentration of SDS and/or increasing the temperature of the hybridization and/or wash. Those skilled in the art will be aware that the conditions for hybridization and/or wash may vary depending upon the nature of the hybridization matrix used to support the sample RNA, or the type of hybridization probe used.

In general, the sample or the probe is immobilized on a solid matrix or surface (e.g., nitrocellulose). For high throughput screening, the sample or probe will generally comprise an array of nucleic acids on glass or other solid matrix, such as, for example, as described in WO 96/17958. Techniques for producing high density arrays are described, for example, by Fodor *et al.*, Science 767-773, 1991, and in U.S. Pat. No. 5,143,854. Typical protocols for other assay formats can be found, for example in Current Protocols In Molecular Biology, Unit 2 (Northern Blotting), Unit 4 (Southern Blotting), and Unit 18 (PCR Analysis), Frederick M. Ausubul *et al.* (ed)., 1995.

10 The detection means according to this aspect of the invention may be any nucleic acid-based detection means such as, for example, nucleic acid hybridization or amplification reaction (eg. PCR), a nucleic acid sequence-based amplification (NASBA) system, inverse polymerase chain reaction (IPCR), *in situ* polymerase chain reaction, or reverse transcription polymerase chain reaction (RT-PCR), amongst others.

15 The probe can be labelled with a reporter molecule capable of producing an identifiable signal (e.g., a radioisotope such as  $^{32}\text{P}$  or  $^{35}\text{S}$ , or a fluorescent or biotinylated molecule). According to this embodiment, those skilled in the art will be aware that the detection of said reporter molecule provides for identification of the probe and that, following the hybridization reaction, the detection of the corresponding nucleotide sequences in the sample is facilitated. Additional probes can be used to confirm the assay results obtained using a single probe.

25 Wherein the detection means is an amplification reaction such as, for example, a polymerase chain reaction or a nucleic acid sequence-based amplification (NASBA) system or a variant thereof, one or more nucleic acid probes molecules of at least about 20 contiguous nucleotides in length is hybridized to mRNA encoding a cancer-associated protein, or alternatively, hybridized to cDNA or cRNA produced from said mRNA, and nucleic acid copies of the template are enzymically-amplified.

30 Those skilled in the art will be aware that there must be a sufficiently high percentage of nucleotide sequence identity between the probes and the RNA sequences in the sample template molecule for hybridization to occur. As stated previously, the stringency conditions can be selected to promote hybridization.

35

In one format, PCR provides for the hybridization of non-complementary probes to different strands of a double-stranded nucleic acid template molecule (ie. a DNA/RNA, RNA/RNA or DNA/DNA template), such that the hybridized probes are positioned to facilitate the 5'-to 3' synthesis of nucleic acid in the intervening region, under the control of a thermostable DNA polymerase enzyme. In accordance with this embodiment, one sense probe and one antisense probe as described herein would be used to amplify DNA from the hybrid RNA/DNA template or cDNA.

In the present context, the cDNA would generally be produced by reverse transcription of mRNA present in the sample being tested (ie. RT-PCR). RT-PCR is particularly useful when it is desirable to determine expression of a cancer-associated gene. It is also known to those skilled in the art to use mRNA/DNA hybrid molecules as a template for such amplification reactions, and, as a consequence, first strand cDNA synthesis is all that is required to be performed prior to the amplification reaction.

Variations of the embodiments described herein are described in detail by McPherson *et al.*, PCR: A Practical Approach. (series eds, D. Rickwood and B.D. Hames), IRL Press Limited, Oxford. pp1-253, 1991.

The amplification reaction detection means described *supra* can be further coupled to a classical hybridization reaction detection means to further enhance sensitivity and specificity of the inventive method, such as by hybridizing the amplified DNA with a probe which is different from any of the probes used in the amplification reaction.

Similarly, the hybridization reaction detection means described *supra* can be further coupled to a second hybridization step employing a probe which is different from the probe used in the first hybridization reaction.

The comparison to be performed in accordance with the present invention may be a visual comparison of the signal generated by the probe, or alternatively, a comparison of data integrated from the signal, such as, for example, data that have been corrected or normalized to allow for variation between samples. Such comparisons can be readily performed by those skilled in the art.

### *Polypeptides*

Cancer-associated polypeptides are encoded by cancer-associated genes. It will be understood that such polypeptides include those polypeptide and fragments thereof that  
5 are homologous to the polypeptides encoded by the nucleotide sequences referred to in Tables 1-3, which are obtained from any source, for example related viral/bacterial proteins, cellular homologues and synthetic peptides, as well as variants or derivatives thereof.

10 Thus, the present invention encompasses the use of variants, homologues or derivatives of the cancer-associated proteins described in the accompanying Tables. In one embodiment, homologues are naturally occurring sequences, such as orthologues, tissue-specific isoforms and allelic variants.

15 In the context of the present invention, a homologous sequence is taken to include an amino acid sequence which is at least 60, 70, 80 or 90% identical, preferably at least 95 or 98% identical at the amino acid level over at least 20, 40, 60 or 80 amino acids with a sequence encoded by a nucleotide sequence referred to in any one of Tables 1-3. In particular, homology should typically be considered with respect to those regions of the  
20 sequence known to be essential for specific biological functions rather than non-essential neighbouring sequences.

Although amino acid homology can also be considered in terms of similarity (i.e. amino acid residues having similar chemical properties/functions), in the context of the present  
25 invention it is preferred to express homology in terms of sequence identity.

Homology comparisons are carried out as described above for nucleotide sequences with the appropriate modifications for amino acid sequences. For example when using the GCG Wisconsin Bestfit package (see below) the default gap penalty for amino acid sequences is  
30 -12 for a gap and -4 for each extension.

It should also be noted that where computer algorithms are used to align amino acid sequences, although the final % homology are measured in terms of identity, the alignment process itself is typically not based on an all-or-nothing pair comparison.  
35 Instead, a scaled similarity score matrix is generally used that assigns scores to each pairwise comparison based on chemical similarity or evolutionary distance. An example

of such a matrix commonly used is the BLOSUM62 matrix - the default matrix for the BLAST suite of programs. GCG Wisconsin programs generally use either the public default values or a custom symbol comparison table if supplied (see user manual for further details). It is preferred to use the public default values for the GCG package, or in  
5 the case of other software, the default matrix, such as BLOSUM62.

The terms "variant" or "derivative" in relation to the amino acid sequences of the present invention includes any substitution of, variation of, modification of, replacement of, deletion  
10 of or addition of one (or more) amino acids from or to the sequence providing the resultant amino acid sequence preferably has biological activity, preferably having at least 25 to 50% of the activity as the polypeptides referred to in the sequence listings, more preferably at least substantially the same activity. Particular details of biological activity for each polypeptide are given in Tables 1-3.

15 Thus, the polypeptides referred to in Tables 1-3 and homologues thereof, are modified for use in the present invention. Typically, modifications are made that maintain the activity of the sequence. Thus, in one embodiment, amino acid substitutions are made, for example from 1, 2 or 3 to 10, 20 or 30 substitutions provided that the modified sequence retains at least about 25 to 50% of, or substantially the same activity.  
20 However, in an alternative preferred embodiment, modifications to the amino acid sequences of a cancer-associated protein are made intentionally to reduce the biological activity of the polypeptide. For example truncated polypeptides that remain capable of binding to target molecules but lack functional effector domains are useful as inhibitors of the biological activity of the full length molecule.

25 In general, preferably less than 20%, 10% or 5% of the amino acid residues of a variant or derivative are altered as compared with the corresponding region of the polypeptides referred to in Tables 1-3.

30 Amino acid substitutions may include the use of non-naturally occurring analogues, for example to increase blood plasma half-life of a therapeutically administered polypeptide (see below for further details on the production of peptide derivatives for use in therapy).

Conservative substitutions are made, for example according to the Table below. Amino  
35 acids in the same block in the second column and preferably in the same line in the third column are substituted for each other:

ALIPHATIC	Non-polar	GAP
		ILV
	Polar - uncharged	CSTM
		NQ
	Polar - charged	DE
		KR
AROMATIC		HFVY

Cancer-associated proteins also include fragments of the above mentioned full length polypeptides and variants thereof, including fragments of the sequences referred to in  
 5 Tables 1-3 and homologues thereof. Preferred fragments include those which include an epitope. Suitable fragments will be at least about 6 or 8, e.g. at least 10, 12, 15 or 20 amino acids in length. They may also be less than 200, 100 or 50 amino acids in length. Polypeptide fragments may contain one or more (e.g. 2, 3, 5, or 10) substitutions, deletions or insertions, including conserved substitutions. Where substitutions, deletion and/or  
 10 insertions have been made, for example by means of recombinant technology, preferably less than 20%, 10% or 5% of the amino acid residues depicted in the sequence listings are altered.

Cancer-associated proteins are preferably in a substantially isolated form. It will be  
 15 understood that the protein are mixed with carriers or diluents which will not interfere with the intended purpose of the protein and still be regarded as substantially isolated. A cancer-associated protein of the invention may also be in a substantially purified form, in which case it will generally comprise the protein in a preparation in which more than 90%, e.g. 95%, 98% or 99% pure as determined by SDS/PAGE or other art-recognized  
 20 means for assessing protein purity.

#### *Protein Production*

For producing full-length polypeptides or immunologically active derivatives thereof by recombinant means, a protein-encoding region comprising at least about 15 contiguous  
 25 nucleotides of the protein-encoding region of a nucleic acid referred to in any one of Tables 1-3 is placed in operable connection with a promoter or other regulatory sequence capable of regulating expression in a cell-free system or cellular system.

Reference herein to a "promoter" is to be taken in its broadest context and includes the transcriptional regulatory sequences of a classical genomic gene, including the TATA box which is required for accurate transcription initiation, with or without a CCAAT box sequence and additional regulatory elements (i.e., upstream activating sequences, enhancers and silencers) which alter gene expression in response to developmental and/or external stimuli, or in a tissue-specific manner. In the present context, the term "promoter" is also used to describe a recombinant, synthetic or fusion molecule, or derivative which confers, activates or enhances the expression of a nucleic acid molecule to which it is operably connected, and which encodes the polypeptide or peptide fragment. Preferred promoters can contain additional copies of one or more specific regulatory elements to further enhance expression and/or to alter the spatial expression and/or temporal expression of the said nucleic acid molecule.

Placing a nucleic acid molecule under the regulatory control of, i.e., "in operable connection with", a promoter sequence means positioning said molecule such that expression is controlled by the promoter sequence. Promoters are generally positioned 5' (upstream) to the coding sequence that they control. To construct heterologous promoter/structural gene combinations, it is generally preferred to position the promoter at a distance from the gene transcription start site that is approximately the same as the distance between that promoter and the gene it controls in its natural setting, i.e., the gene from which the promoter is derived. Furthermore, the regulatory elements comprising a promoter are usually positioned within 2 kb of the start site of transcription of the gene. As is known in the art, some variation in this distance can be accommodated without loss of promoter function. Similarly, the preferred positioning of a regulatory sequence element with respect to a heterologous gene to be placed under its control is defined by the positioning of the element in its natural setting, i.e., the genes from which it is derived. Again, as is known in the art, some variation in this distance can also occur.

The prerequisite for producing intact polypeptides and peptides in bacteria such as *E. coli* is the use of a strong promoter with an effective ribosome binding site. Typical promoters suitable for expression in bacterial cells such as *E. coli* include, but are not limited to, the *lacZ* promoter, temperature-sensitive  $\lambda_L$  or  $\lambda_R$  promoters, T7 promoter or the IPTG-inducible *tac* promoter. A number of other vector systems for expressing the nucleic acid molecule of the invention in *E. coli* are well-known in the art and are described, for example, in Ausubel *et al* (*In: Current Protocols in Molecular Biology*. Wiley Interscience, ISBN 047150338, 1987) or Sambrook *et al* (*In: Molecular cloning*. A

laboratory manual, second edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1989). Numerous plasmids with suitable promoter sequences for expression in bacteria and efficient ribosome binding sites have been described, such as for example, pKC30 ( $\lambda$ : Shimatake and Rosenberg, *Nature* 292, 128, 1981); pKK173-3 (*tac*: Amann and Brosius, *Gene* 40, 183, 1985), pET-3 (T7: Studier and Moffat, *J. Mol. Biol.* 189, 113, 1986); the pBAD/TOPO or pBAD/Thio-TOPO series of vectors containing an arabinose-inducible promoter (Invitrogen, Carlsbad, CA), the latter of which is designed to also produce fusion proteins with thioredoxin to enhance solubility of the expressed protein; the pFLEX series of expression vectors (Pfizer Inc., CT, USA); or the pQE series of expression vectors (Qiagen, CA), amongst others.

Typical promoters suitable for expression in viruses of eukaryotic cells and eukaryotic cells include the SV40 late promoter, SV40 early promoter and cytomegalovirus (CMV) promoter, CMV IE (cytomegalovirus immediate early) promoter amongst others. Preferred vectors for expression in mammalian cells (eg. 293, COS, CHO, 293T cells) include, but are not limited to, the pcDNA vector suite supplied by Invitrogen, in particular pcDNA 3.1 myc-His-tag comprising the CMV promoter and encoding a C-terminal 6xHis and MYC tag; and the retrovirus vector pSRatkneo (Muller *et al.*, *Mol. Cell. Biol.*, 11, 1785, 1991). The vector pcDNA 3.1 myc-His (Invitrogen) is particularly preferred for expressing a secreted form of a protein in 293T cells, wherein the expressed peptide or protein can be purified free of conspecific proteins, using standard affinity techniques that employ a Nickel column to bind the protein via the His tag.

A wide range of additional host/vector systems suitable for expressing polypeptides or immunological derivatives thereof are available publicly, and described, for example, in Sambrook *et al* (*In*: Molecular cloning. A laboratory manual, second edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1989).

Means for introducing the isolated nucleic acid molecule or a gene construct comprising same into a cell for expression are well-known to those skilled in the art. The technique used for a given organism depends on the known successful techniques. Means for introducing recombinant DNA into animal cells include microinjection, transfection mediated by DEAE-dextran, transfection mediated by liposomes such as by using lipofectamine (Gibco, MD, USA) and/or cellfectin (Gibco, MD, USA), PEG-mediated DNA uptake, electroporation and microparticle bombardment such as by using DNA-coated tungsten or gold particles (Agracetus Inc., WI, USA) amongst others.



For producing mutants, nucleotide insertion derivatives of the protein-encoding region are produced by making 5' and 3' terminal fusions, or by making intra-sequence insertions of single or multiple nucleotides or nucleotide analogues. Insertion nucleotide  
5 sequence variants are produced by introducing one or more nucleotides or nucleotide analogues into a predetermined site in the nucleotide sequence of said sequence, although random insertion is also possible with suitable screening of the resulting product being performed. Deletion variants are produced by removing one or more nucleotides from the nucleotide sequence. Substitutional nucleotide variants are produced by  
10 substituting at least one nucleotide in the sequence with a different nucleotide or a nucleotide analogue in its place, with the immunologically active derivative encoded therefor having an identical amino acid sequence, or only a limited number of amino acid modifications that do not alter its antigenicity compared to the base peptide or its ability to bind antibodies prepared against the base peptide. Such mutant derivatives will  
15 preferably have at least 80% identity with the base amino acid sequence from which they are derived.

Preferred immunologically active derivatives of a full-length polypeptide encoded by a gene referred to in any one of Tables 1-3 will comprise at least about 5-10 contiguous  
20 amino acids of the full-length amino acid sequence, more preferably at least about 10-20 contiguous amino acids in length, and even more preferably 20-30 contiguous amino acids in length.

For the purposes of producing derivatives using standard peptide synthesis techniques, such as, for example, Fmoc chemistry, a length not exceeding about 30-50 amino acids  
25 in length is preferred, as longer peptides are difficult to produce at high efficiency. Longer peptide fragments are readily achieved using recombinant DNA techniques wherein the peptide is expressed in a cell-free or cellular expression system comprising nucleic acid encoding the desired peptide fragment.

30 It will be apparent to the skilled artisan that any sufficiently antigenic region of at least about 5-10 amino acid residues can be used to prepare antibodies that bind generally to the polypeptides listed in Tables 1-3.

35 An expressed protein or synthetic peptide is preferably produced as a recombinant fusion protein, such as for example, to aid in extraction and purification. To produce a fusion

polypeptide, the open reading frames are covalently linked in the same reading frame, such as, for example, using standard cloning procedures as described by Ausubel *et al.* (Current Protocols in Molecular Biology, Wiley Interscience, ISBN 047150338, 1992), and expressed under control of a promoter. Examples of fusion protein partners include glutathione-S-transferase (GST), FLAG, hexahistidine, GAL4 (DNA binding and/or transcriptional activation domains) and  $\beta$ -galactosidase. It may also be convenient to include a proteolytic cleavage site between the fusion protein partner and the protein sequence of interest to allow removal of fusion protein sequences. Preferably the fusion protein will not hinder the immune function of the target protein.

In a particularly preferred embodiment, polypeptides are produced substantially free of conspecific proteins. Such purity can be assessed by standard procedures, such as, for example, SDS/polyacrylamide gel electrophoresis, 2-dimensional gene electrophoresis, chromatography, amino acid composition analysis, or amino acid sequence analysis.

To produce isolated polypeptides or fragments, eg., for antibody production, standard protein purification techniques may be employed. For example, gel filtration, ion exchange chromatography, reverse phase chromatography, or affinity chromatography, or a combination of any one or more said procedures, may be used. High pressure and low pressure procedures can also be employed, such as, for example, FPLC, or HPLC. To isolate the full-length proteins or peptide fragments comprising more than about 50-100 amino acids in length, it is particularly preferred to express the polypeptide in a suitable cellular expression system in combination with a suitable affinity tag, such as a 6xHis tag, and to purify the polypeptide using an affinity step that bonds it via the tag (*supra*). Optionally, the tag may then be cleaved from the expressed polypeptide.

Alternatively, for short immunologically active derivatives of a full-length polypeptide, preferably those peptide fragments comprising less than about 50 amino acids in length, chemical synthesis techniques are conveniently used. As will be known to those skilled in the art, such techniques may also produce contaminating peptides that are shorter than the desired peptide, in which case the desired peptide is conveniently purified using reverse phase and/or ion exchange chromatography procedures at high pressure (ie. HPLC or FPLC).

### Antibodies

The invention also provides monoclonal or polyclonal antibodies that bind specifically to polypeptides of the invention or fragments thereof. Thus, the present invention further provides a process for the production of monoclonal or polyclonal antibodies to polypeptides of the invention.

5

The phrase "binds specifically" to a polypeptide means that the binding of the antibody to the protein or peptide is determinative of the presence of the protein, in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein at least two times the background and more typically more than 10 to 100 times background. Typically, antibodies of the invention bind to a protein of interest with a  $K_d$  of at least about 0.1 mM, more usually at least about 1  $\mu$ M, preferably at least about 0.1  $\mu$ M, and most preferably at least, 0.01  $\mu$ M.

15 Reference herein to antibody or antibodies includes whole polyclonal and monoclonal antibodies, and parts thereof, either alone or conjugated with other moieties. Antibody parts include Fab and  $F(ab)_2$  fragments and single chain antibodies. The antibodies may be made *in vivo* in suitable laboratory animals, or, in the case of engineered antibodies (Single Chain Antibodies or SCABS, etc) using recombinant DNA techniques *in vitro*.

20

In accordance with this aspect of the invention, the antibodies may be produced for the purposes of immunizing the subject, in which case high titer or neutralizing antibodies that bind to a B cell epitope will be especially preferred. Suitable subjects for immunization will, of course, depend upon the immunizing antigen or antigenic B cell epitope. It is contemplated that the present invention will be broadly applicable to the immunization of a wide range of animals, such as, for example, farm animals (e.g. horses, cattle, sheep, pigs, goats, chickens, ducks, turkeys, and the like), laboratory animals (e.g. rats, mice, guinea pigs, rabbits), domestic animals (cats, dogs, birds and the like), feral or wild exotic animals (e.g. possums, cats, pigs, buffalo, wild dogs and the like) and humans.

30

Alternatively, the antibodies may be for commercial or diagnostic purposes, in which case the subject to whom the diagnostic/prognostic protein or immunogenic fragment or epitope thereof is administered will most likely be a laboratory or farm animal. A wide range of animal species are used for the production of antisera. Typically the animal used for production of antisera is a rabbit, a mouse, rat, hamster, guinea pig, goat,

35

sheep, pig, dog, horse, or chicken. Because of the relatively large blood volume of rabbits, a rabbit is a preferred choice for production of polyclonal antibodies. However, as will be known to those skilled in the art, larger amounts of immunogen are required to obtain high antibodies from large animals as opposed to smaller animals such as mice.

5 In such cases, it will be desirable to isolate the antibody from the immunized animal.

Preferably, the antibody is a high titer antibody. By "high titer" means a sufficiently high titer to be suitable for use in diagnostic or therapeutic applications. As will be known in the art, there is some variation in what might be considered "high titer". For most

10 applications a titer of at least about  $10^3$ - $10^4$  is preferred. More preferably, the antibody titer will be in the range from about  $10^4$  to about  $10^5$ , even more preferably in the range from about  $10^5$  to about  $10^6$ .

More preferably, in the case of B cell epitopes from pathogens, viruses or bacteria, the

15 antibody is a neutralizing antibody (i.e. it is capable of neutralizing the infectivity of the organism from which the B cell epitope is derived).

To generate antibodies, the diagnostic/prognostic protein or immunogenic fragment or epitope thereof, optionally formulated with any suitable or desired carrier, adjuvant, BRM,

20 or pharmaceutically acceptable excipient, is conveniently administered in the form of an injectable composition. Injection may be intranasal, intramuscular, sub-cutaneous, intravenous, intradermal, intraperitoneal, or by other known route. For intravenous injection, it is desirable to include one or more fluid and nutrient replenishers. Means for preparing and characterizing antibodies are well known in the art. (See, e.g.,

25 ANTIBODIES: A LABORATORY MANUAL, Cold Spring Harbor Laboratory, 1988, incorporated herein by reference).

The efficacy of the diagnostic/prognostic protein or immunogenic fragment or epitope thereof in producing an antibody is established by injecting an animal, for example, a

30 mouse, rat, rabbit, guinea pig, dog, horse, cow, goat or pig, with a formulation comprising the diagnostic/prognostic protein or immunogenic fragment or epitope thereof, and then monitoring the immune response to the B cell epitope, as described in the Examples. Both primary and secondary immune responses are monitored. The antibody titer is determined using any conventional immunoassay, such as, for example, ELISA, or radio

35 immunoassay.

The production of polyclonal antibodies may be monitored by sampling blood of the immunized animal at various points following immunization. A second, booster injection, may be given, if required to achieve a desired antibody titer. The process of boosting and titering is repeated until a suitable titer is achieved. When a desired level of immunogenicity is obtained, the immunized animal is bled and the serum isolated and stored, and/or the animal is used to generate monoclonal antibodies (Mabs).

For the production of monoclonal antibodies (Mabs) any one of a number of well-known techniques may be used, such as, for example, the procedure exemplified in US Patent No. 4,196,265, incorporated herein by reference.

For example, a suitable animal will be immunized with an effective amount of the diagnostic/prognostic protein or immunogenic fragment or epitope thereof under conditions sufficient to stimulate antibody producing cells. Rodents such as mice and rats are preferred animals, however, the use of rabbit, sheep, or frog cells is also possible. The use of rats may provide certain advantages, but mice are preferred, with the BALB/c mouse being most preferred as the most routinely used animal and one that generally gives a higher percentage of stable fusions.

Following immunization, somatic cells with the potential for producing antibodies, specifically B lymphocytes (B cells), are selected for use in the MAb generating protocol. These cells may be obtained from biopsied spleens, tonsils or lymph nodes, or from a peripheral blood sample. Spleen cells and peripheral blood cells are preferred, the former because they are a rich source of antibody-producing cells that are in the dividing plasmablast stage, and the latter because peripheral blood is easily accessible. Often, a panel of animals will have been immunized and the spleen of animal with the highest antibody titer removed. Spleen lymphocytes are obtained by homogenizing the spleen with a syringe. Typically, a spleen from an immunized mouse contains approximately  $5 \times 10^7$  to  $2 \times 10^8$  lymphocytes.

The B cells from the immunized animal are then fused with cells of an immortal myeloma cell, generally derived from the same species as the animal that was immunized with the diagnostic/prognostic protein or immunogenic fragment or epitope thereof. Myeloma cell lines suited for use in hybridoma-producing fusion procedures preferably are non-antibody-producing, have high fusion efficiency and enzyme deficiencies that render them incapable of growing in certain selective media which support the growth of only the

desired fused cells, or hybridomas. Any one of a number of myeloma cells may be used and these are known to those of skill in the art (e.g. murine P3-X63/Ag8, X63-Ag8.653, NS1/1.Ag 4 1, Sp210-Ag14, FO, NSO/U, MPC-11, MPC11-X45-GTG 1.7 and S194/5XX0; or rat R210.RCY3, Y3-Ag 1.2.3, IR983F and 4B210; and U-266, GM1500-GRG2, LICR-LON-HMy2 and UC729-6). A preferred murine myeloma cell is the NS-1 myeloma cell line (also termed P3-NS-1-Ag4-1), which is readily available from the NIGMS Human Genetic Mutant Cell Repository under Accession No. GM3573. Alternatively, a murine myeloma SP2/0 non-producer cell line that is 8-azaguanine-resistant is used.

10

To generate hybrids of antibody-producing spleen or lymph node cells and myeloma cells, somatic cells are mixed with myeloma cells in a proportion between about 20:1 to about 1:1, respectively, in the presence of an agent or agents (chemical or electrical) that promote the fusion of cell membranes. Fusion methods using Sendai virus have been described by Kohler and Milstein, *Nature* 256, 495-497, 1975; and Kohler and Milstein, *Eur. J. Immunol.* 6, 511-519, 1976. Methods using polyethylene glycol (PEG), such as 37% (v/v) PEG, are described in detail by Gefter *et al.*, *Somatic Cell Genet.* 3, 231-236, 1977. The use of electrically induced fusion methods is also appropriate.

Hybrids are amplified by culture in a selective medium comprising an agent that blocks the *de novo* synthesis of nucleotides in the tissue culture media. Exemplary and preferred agents are aminopterin, methotrexate and azaserine. Aminopterin and methotrexate block *de novo* synthesis of both purines and pyrimidines, whereas azaserine blocks only purine synthesis. Where aminopterin or methotrexate is used, the media is supplemented with hypoxanthine and thymidine as a source of nucleotides (HAT medium). Where azaserine is used, the media is supplemented with hypoxanthine.

The preferred selection medium is HAT, because only those hybridomas capable of operating nucleotide salvage pathways are able to survive in HAT medium, whereas myeloma cells are defective in key enzymes of the salvage pathway, (e.g., hypoxanthine phosphoribosyl transferase or HPRT), and they cannot survive. B cells can operate this salvage pathway, but they have a limited life span in culture and generally die within about two weeks. Accordingly, the only cells that can survive in the selective media are those hybrids formed from myeloma and B cells.

35

The amplified hybridomas are subjected to a functional selection for antibody specificity and/or titer, such as, for example, by immunoassay (e.g. radioimmunoassay, enzyme immunoassay, cytotoxicity assay, plaque assay, dot immunobinding assay, and the like).

- 5 The selected hybridomas are serially diluted and cloned into individual antibody-producing cell lines, which clones can then be propagated indefinitely to provide MABs. The cell lines may be exploited for MAb production in two basic ways. A sample of the hybridoma is injected, usually in the peritoneal cavity, into a histocompatible animal of the type that was used to provide the somatic and myeloma cells for the original fusion.
- 10 The injected animal develops tumors secreting the specific monoclonal antibody produced by the fused cell hybrid. The body fluids of the animal, such as serum or ascites fluid, can then be tapped to provide MABs in high concentration. The individual cell lines could also be cultured *in vitro*, where the MABs are naturally secreted into the culture medium from which they are readily obtained in high concentrations. MABs
- 15 produced by either means may be further purified, if desired, using filtration, centrifugation and various chromatographic methods such as HPLC or affinity chromatography.

- Monoclonal antibodies of the present invention also include anti-idiotypic antibodies
- 20 produced by methods well-known in the art. Monoclonal antibodies according to the present invention also may be monoclonal heteroconjugates, (i.e., hybrids of two or more antibody molecules). In another embodiment, monoclonal antibodies according to the invention are chimeric monoclonal antibodies. In one approach, the chimeric monoclonal antibody is engineered by cloning recombinant DNA containing the promoter, leader, and
  - 25 variable-region sequences from a mouse anti-PSA producing cell and the constant-region exons from a human antibody gene. The antibody encoded by such a recombinant gene is a mouse-human chimera. Its antibody specificity is determined by the variable region derived from mouse sequences. Its isotype, which is determined by the constant region, is derived from human DNA.

- 30 In another embodiment, the monoclonal antibody according to the present invention is a "humanized" monoclonal antibody, produced by any one of a number of techniques well-known in the art. That is, mouse complementary determining regions ("CDRs") are transferred from heavy and light V-chains of the mouse Ig into a human V-domain,
- 35 followed by the replacement of some human residues in the framework regions of their

murine counterparts. "Humanized" monoclonal antibodies in accordance with this invention are especially suitable for use *in vivo* in diagnostic and therapeutic methods.

5 As stated above, the monoclonal antibodies and fragments thereof according to this invention are multiplied according to *in vitro* and *in vivo* methods well-known in the art. Multiplication *in vitro* is carried out in suitable culture media such as Dulbecco's modified Eagle medium or RPMI 1640 medium, optionally replenished by a mammalian serum such as fetal calf serum or trace elements and growth-sustaining supplements, e.g.,  
10 feeder cells, such as normal mouse peritoneal exudate cells, spleen cells, bone marrow macrophages or the like. *In vitro* production provides relatively pure antibody preparations and allows scale-up to give large amounts of the desired antibodies. Techniques for large scale hybridoma cultivation under tissue culture conditions are known in the art and include homogenous suspension culture, (e.g., in an airlift reactor or in a continuous stirrer reactor or immobilized or entrapped cell culture).

15 Large amounts of the monoclonal antibody of the present invention also may be obtained by multiplying hybridoma cells *in vivo*. Cell clones are injected into mammals which are histocompatible with the parent cells, (e.g., syngeneic mice, to cause growth of antibody-producing tumors. Optionally, the animals are primed with a hydrocarbon, especially oils  
20 such as Pristane (tetramethylpentadecane) prior to injection.

In accordance with the present invention, fragments of the monoclonal antibody of the invention are obtained from monoclonal antibodies produced as described above, by methods which include digestion with enzymes such as pepsin or papain and/or  
25 cleavage of disulfide bonds by chemical reduction. Alternatively, monoclonal antibody fragments encompassed by the present invention are synthesized using an automated peptide synthesizer, or they may be produced manually using techniques well known in the art.

30 The monoclonal conjugates of the present invention are prepared by methods known in the art, e.g., by reacting a monoclonal antibody prepared as described above with, for instance, an enzyme in the presence of a coupling agent such as glutaraldehyde or periodate. Conjugates with fluorescein markers are prepared in the presence of these coupling agents, or by reaction with an isothiocyanate. Conjugates with metal chelates  
35 are similarly produced. Other moieties to which antibodies may be conjugated include



radionuclides such as, for example,  $^3\text{H}$ ,  $^{125}\text{I}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^{14}\text{C}$ ,  $^{51}\text{Cr}$ ,  $^{36}\text{Cl}$ ,  $^{57}\text{Co}$ ,  $^{58}\text{Co}$ ,  $^{59}\text{Fe}$ ,  $^{75}\text{Se}$ , and  $^{152}\text{Eu}$ .

Radioactively labeled monoclonal antibodies of the present invention are produced according to well-known methods in the art. For instance, monoclonal antibodies are iodinated by contact with sodium or potassium iodide and a chemical oxidizing agent such as sodium hypochlorite, or an enzymatic oxidizing agent, such as lactoperoxidase. Monoclonal antibodies according to the invention may be labeled with technetium<sup>99</sup> by ligand exchange process, for example, by reducing pertechnetate with stannous solution, chelating the reduced technetium onto a Sephadex column and applying the antibody to this column or by direct labeling techniques, (e.g., by incubating pertechnetate, a reducing agent such as  $\text{SNCl}_2$ , a buffer solution such as sodium-potassium phthalate solution, and the antibody).

Any immunoassay may be used to monitor antibody production by the diagnostic/prognostic protein or immunogenic fragment or epitope thereof. Immunoassays, in their most simple and direct sense, are binding assays. Certain preferred immunoassays are the various types of enzyme linked immunosorbent assays (ELISAs) and radioimmunoassays (RIA) known in the art. Immunohistochemical detection using tissue sections is also particularly useful. However, it will be readily appreciated that detection is not limited to such techniques, and Western blotting, dot blotting, FACS analyses, and the like may also be used.

Most preferably, the assay will be capable of generating quantitative results.

25

For example, antibodies are tested in simple competition assays. A known antibody preparation that binds to the B cell epitope and the test antibody are incubated with an antigen composition comprising the B cell epitope, preferably in the context of the native antigen. "Antigen composition" as used herein means any composition that contains some version of the B cell epitope in an accessible form. Antigen-coated wells of an ELISA plate are particularly preferred. In one embodiment, one would pre-mix the known antibodies with varying amounts of the test antibodies (e.g., 1:1, 1:10 and 1:100) for a period of time prior to applying to the antigen composition. If one of the known antibodies is labeled, direct detection of the label bound to the antigen is possible; comparison to an unmixed sample assay will determine competition by the test antibody and, hence, cross-

35

reactivity. Alternatively, using secondary antibodies specific for either the known or test antibody, one will be able to determine competition.

An antibody that binds to the antigen composition will be able to effectively compete for  
5 binding of the known antibody and thus will significantly reduce binding of the latter. The reactivity of the known antibodies in the absence of any test antibody is the control. A significant reduction in reactivity in the presence of a test antibody is indicative of a test antibody that binds to the B cell epitope (i.e., it cross-reacts with the known antibody).

- 10 In one exemplary ELISA, the antibodies against the diagnostic/prognostic protein or immunogenic fragment or B cell epitope are immobilized onto a selected surface exhibiting protein affinity, such as a well in a polystyrene microtiter plate. Then, a composition containing a peptide comprising the B cell epitope is added to the wells. After binding and washing to remove non-specifically bound immune complexes, the  
15 bound epitope may be detected. Detection is generally achieved by the addition of a second antibody that is known to bind to the B cell epitope and is linked to a detectable label. This type of ELISA is a simple "sandwich ELISA". Detection may also be achieved by the addition of said second antibody, followed by the addition of a third antibody that has binding affinity for the second antibody, with the third antibody being linked to a  
20 detectable label.

Antibodies of the invention may be bound to a solid support and/or packaged into kits in a suitable container along with suitable reagents, controls, instructions and the like.

#### 25 *Immunoassay formats*

- In one embodiment, a cancer-associated protein or an immunogenic fragment or epitope thereof is detected in a patient sample, wherein the level of the protein or immunogenic fragment or epitope in the sample is indicative of ovarian cancer or disease recurrence or an indicator of poor survival. Preferably, the method comprises contacting a biological  
30 sample derived from the subject with an antibody capable of binding to a cancer-associated protein or an immunogenic fragment or epitope thereof, and detecting the formation of an antigen-antibody complex.

- In another embodiment, an antibody against a cancer-associated protein or epitope  
35 thereof is detected in a patient sample, wherein the level of the antibody in the sample is indicative of ovarian cancer or disease recurrence or an indicator of poor survival.

Preferably, the method comprises contacting a biological sample derived from the subject with a cancer-associated protein or an antigenic fragment eg., a B cell epitope or other immunogenic fragment thereof, and detecting the formation of an antigen-antibody complex.

5

The diagnostic assays of the invention are useful for determining the progression of ovarian cancer or a metastasis thereof in a subject. In accordance with these prognostic applications of the invention, the level of a cancer-associated protein or an immunogenic fragment or epitope thereof in a biological sample is correlated with the disease state eg., as determined by clinical symptoms or biochemical tests (eg., CA125 levels).

10

Accordingly, a further embodiment of the invention provides a method for detecting a cancer cell in a subject, said method comprising:

- (i) determining the level of a cancer-associate protein in a test sample from said subject; and
  - (ii) comparing the level determined at (i) to the level of said cancer-associated protein in a comparable sample from a healthy or normal individual,
- wherein a level of said cancer-associate protein at (i) that is modified in the test sample relative to the comparable sample from the normal or healthy individual is indicative of the presence of a cancer cell in said subject.

15

20

In one embodiment of the diagnostic/prognostic methods described herein, the biological sample is obtained previously from the subject. In accordance with such an embodiment, the prognostic or diagnostic method is performed *ex vivo*.

25

In yet another embodiment, the subject diagnostic/prognostic methods further comprise processing the sample from the subject to produce a derivative or extract that comprises the analyte.

Preferred detection systems contemplated herein include any known assay for detecting proteins or antibodies in a biological sample isolated from a human subject, such as, for example, SDS/PAGE, isoelectric focussing, 2-dimensional gel electrophoresis comprising SDS/PAGE and isoelectric focussing, an immunoassay, a detection based system using an antibody or non-antibody ligand of the protein, such as, for example, a small molecule (e.g. a chemical compound, agonist, antagonist, allosteric modulator, competitive inhibitor, or non-competitive inhibitor, of the protein). In accordance with

30

35

these embodiments, the antibody or small molecule may be used in any standard solid phase or solution phase assay format amenable to the detection of proteins. Optical or fluorescent detection, such as, for example, using mass spectrometry, MALDI-TOF, biosensor technology, evanescent fiber optics, or fluorescence resonance energy transfer, is clearly encompassed by the present invention. Assay systems suitable for use in high throughput screening of mass samples, particularly a high throughput spectroscopy resonance method (e.g. MALDI-TOF, electrospray MS or nano-electrospray MS), are particularly contemplated.

- 10 Immunoassay formats are particularly preferred, eg., selected from the group consisting of, an immunoblot, a Western blot, a dot blot, an enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA), enzyme immunoassay. Modified immunoassays utilizing fluorescence resonance energy transfer (FRET), isotope-coded affinity tags (ICAT), matrix-assisted laser desorption/ionization time of flight (MALDI-TOF),  
15 electrospray ionization (ESI), biosensor technology, evanescent fiber-optics technology or protein chip technology are also useful.

Preferably, the assay is a semi-quantitative assay or quantitative assay.

- 20 Standard solid phase ELISA formats are particularly useful in determining the concentration of a protein or antibody from a variety of patient samples.

In one form such as an assay involves immobilising a biological sample comprising antibodies against the cancer-associated protein or epitope, or alternatively an ovarian  
25 cancer-associated protein or an immunogenic fragment thereof, onto a solid matrix, such as, for example a polystyrene or polycarbonate microwell or dipstick, a membrane, or a glass support (e.g. a glass slide).

In the case of an antigen-based assay, an antibody that specifically binds an ovarian  
30 cancer-associated protein is brought into direct contact with the immobilised biological sample, and forms a direct bond with any of its target protein present in said sample. For an antibody-based assay, an immobilized ovarian cancer-associated protein or an immunogenic fragment or epitope thereof is contacted with the sample. The added antibody or protein in solution is generally labelled with a detectable reporter molecule,  
35 such as for example, a fluorescent label (e.g. FITC or Texas Red) or an enzyme (e.g. horseradish peroxidase (HRP)), alkaline phosphatase (AP) or  $\beta$ -galactosidase.

Alternatively, or in addition, a second labelled antibody can be used that binds to the first antibody or to the isolated/recombinant antigen. Following washing to remove any unbound antibody or antigen, as appropriate, the label is detected either directly, in the case of a fluorescent label, or through the addition of a substrate, such as for example hydrogen peroxide, TMB, or toluidine, or 5-bromo-4-chloro-3-indol-beta-D-galactopyranoside (x-gal).

Such ELISA based systems are particularly suitable for quantification of the amount of a protein or antibody in a sample, such as, for example, by calibrating the detection system against known amounts of a standard.

In another form, an ELISA consists of immobilizing an antibody that specifically binds an ovarian cancer-associated protein on a solid matrix, such as, for example, a membrane, a polystyrene or polycarbonate microwell, a polystyrene or polycarbonate dipstick or a glass support. A patient sample is then brought into physical relation with said antibody, and the antigen in the sample is bound or 'captured'. The bound protein can then be detected using a labelled antibody. For example if the protein is captured from a human sample, an anti-human antibody is used to detect the captured protein. Alternatively, a third labelled antibody can be used that binds the second (detecting) antibody.

It will be apparent to the skilled person that the assay formats described herein are amenable to high throughput formats, such as, for example automation of screening processes, or a microarray format as described in Mendoza *et al*, Biotechniques 27(4): 778-788, 1999. Furthermore, variations of the above described assay will be apparent to those skilled in the art, such as, for example, a competitive ELISA.

Alternatively, the presence of antibodies against the cancer-associate protein, or alternatively an ovarian cancer-associated protein or an immunogenic fragment thereof, is detected using a radioimmunoassay (RIA). The basic principle of the assay is the use of a radiolabelled antibody or antigen to detect antibody antigen interactions. For example, an antibody that specifically binds to an ovarian cancer-associated protein can be bound to a solid support and a biological sample brought into direct contact with said antibody. To detect the bound antigen, an isolated and/or recombinant form of the antigen is radiolabelled is brought into contact with the same antibody. Following washing the amount of bound radioactivity is detected. As any antigen in the biological sample inhibits binding of the radiolabelled antigen the amount of radioactivity detected is inversely

proportional to the amount of antigen in the sample. Such an assay may be quantitated by using a standard curve using increasing known concentrations of the isolated antigen.

As will be apparent to the skilled artisan, such an assay may be modified to use any reporter molecule, such as, for example, an enzyme or a fluorescent molecule, in place of a radioactive label.

Western blotting is also useful for detecting an ovarian cancer-associated protein or an immunogenic fragment thereof. In such an assay protein from a biological sample is separated using sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (SDS-PAGE) using techniques well known in the art and described in, for example, Scopes (*In: Protein Purification: Principles and Practice, Third Edition, Springer Verlag, 1994*). Separated proteins are then transferred to a solid support, such as, for example, a membrane or more specifically PVDF membrane, using methods well known in the art, for example, electrotransfer. This membrane may then be blocked and probed with a labelled antibody or ligand that specifically binds an ovarian cancer-associated protein. Alternatively, a labelled secondary, or even tertiary, antibody or ligand can be used to detect the binding of a specific primary antibody.

High-throughput methods for detecting the presence or absence of antibodies, or alternatively ovarian cancer-associated protein or an immunogenic fragment thereof are particularly preferred.

In one embodiment, MALDI-TOF is used for the rapid identification of a protein. Accordingly, there is no need to detect the proteins of interest using an antibody or ligand that specifically binds to the protein of interest. Rather, proteins from a biological sample are separated using gel electrophoresis using methods well known in the art and those proteins at approximately the correct molecular weight and/or isoelectric point are analysed using MALDI-TOF to determine the presence or absence of a protein of interest.

Alternatively, MALDI or ESI or a combination of approaches is used to determine the concentration of a particular protein in a biological sample, such as, for example sputum. Such proteins are preferably well characterised previously with regard to parameters such as molecular weight and isoelectric point.

Biosensor devices generally employ an electrode surface in combination with current or impedance measuring elements to be integrated into a device in combination with the assay substrate (such as that described in U.S. Patent No. 5,567,301). An antibody or ligand that specifically binds to a protein of interest is preferably incorporated onto the surface of a biosensor device and a biological sample isolated from a patient (for example sputum that has been solubilised using the methods described herein) contacted to said device. A change in the detected current or impedance by the biosensor device indicates protein binding to said antibody or ligand. Some forms of biosensors known in the art also rely on surface plasmon resonance to detect protein interactions, whereby a change in the surface plasmon resonance surface of reflection is indicative of a protein binding to a ligand or antibody (U.S. Patent No. 5,485,277 and 5,492,840).

Biosensors are of particular use in high throughput analysis due to the ease of adapting such systems to micro- or nano-scales. Furthermore, such systems are conveniently adapted to incorporate several detection reagents, allowing for multiplexing of diagnostic reagents in a single biosensor unit. This permits the simultaneous detection of several epitopes in a small amount of body fluids.

Evanescent biosensors are also preferred as they do not require the pretreatment of a biological sample prior to detection of a protein of interest. An evanescent biosensor generally relies upon light of a predetermined wavelength interacting with a fluorescent molecule, such as for example, a fluorescent antibody attached near the probe's surface, to emit fluorescence at a different wavelength upon binding of the diagnostic protein to the antibody or ligand.

To produce protein chips, the proteins, peptides, polypeptides, antibodies or ligands that are able to bind specific antibodies or proteins of interest are bound to a solid support such as for example glass, polycarbonate, polytetrafluoroethylene, polystyrene, silicon oxide, metal or silicon nitride. This immobilization is either direct (e.g. by covalent linkage, such as, for example, Schiff's base formation, disulfide linkage, or amide or urea bond formation) or indirect. Methods of generating a protein chip are known in the art and are described in for example U.S. Patent Application No. 20020136821, 20020192654, 20020102617 and U.S. Patent No. 6,391,625. In order to bind a protein to a solid support it is often necessary to treat the solid support so as to create chemically reactive groups on the surface, such as, for example, with an aldehyde-containing silane reagent.

Alternatively, an antibody or ligand may be captured on a microfabricated polyacrylamide gel pad and accelerated into the gel using microelectrophoresis as described in, Arenkov *et al. Anal. Biochem.* 278:123-131, 2000.

- 5 A protein chip is preferably generated such that several proteins, ligands or antibodies are arrayed on said chip. This format permits the simultaneous screening for the presence of several proteins in a sample.

- 10 Alternatively, a protein chip may comprise only one protein, ligand or antibody, and be used to screen one or more patient samples for the presence of one polypeptide of interest. Such a chip may also be used to simultaneously screen an array of patient samples for a polypeptide of interest.

- 15 Preferably, a sample to be analysed using a protein chip is attached to a reporter molecule, such as, for example, a fluorescent molecule, a radioactive molecule, an enzyme, or an antibody that is detectable using methods well known in the art. Accordingly, by contacting a protein chip with a labelled sample and subsequent washing to remove any unbound proteins the presence of a bound protein is detected using methods well known in the art, such as, for example using a DNA microarray reader.

- 20 Alternatively, biomolecular interaction analysis-mass spectrometry (BIA-MS) is used to rapidly detect and characterise a protein present in complex biological samples at the low- to sub-fmole level (Nelson *et al. Electrophoresis* 21: 1155-1163, 2000). One technique useful in the analysis of a protein chip is surface enhanced laser  
25 desorption/ionization-time of flight-mass spectrometry (SELDI-TOF-MS) technology to characterise a protein bound to the protein chip. Alternatively, the protein chip is analysed using ESI as described in U.S. Patent Application 20020139751.

- 30 As will be apparent to the skilled artisan, protein chips are particularly amenable to multiplexing of detection reagents. Accordingly, several antibodies or ligands each able to specifically bind a different peptide or protein may be bound to different regions of said protein chip. Analysis of a biological sample using said chip then permits the detecting of multiple proteins of interest, or multiple B cell epitopes of the ovarian cancer-associated protein. Multiplexing of diagnostic and prognostic markers is particularly contemplated in  
35 the present invention.



In a further embodiment, the samples are analysed using ICAT, essentially as described in US Patent Application No. 20020076739. This system relies upon the labelling of a protein sample from one source (i.e. a healthy individual) with a reagent and the labelling of a protein sample from another source (i.e. a tuberculosis patient) with a second reagent that is chemically identical to the first reagent, but differs in mass due to isotope composition. It is preferable that the first and second reagents also comprise a biotin molecule. Equal concentrations of the two samples are then mixed, and peptides recovered by avidin affinity chromatography. Samples are then analysed using mass spectrometry. Any difference in peak heights between the heavy and light peptide ions directly correlates with a difference in protein abundance in a biological sample. The identity of such proteins may then be determined using a method well known in the art, such as, for example MALDI-TOF, or ESI.

As will be apparent to those skilled in the art a diagnostic or prognostic assay described herein may be a multiplexed assay. As used herein the term "multiplex", shall be understood not only to mean the detection of two or more diagnostic or prognostic markers in a single sample simultaneously, but also to encompass consecutive detection of two or more diagnostic or prognostic markers in a single sample, simultaneous detection of two or more diagnostic or prognostic markers in distinct but matched samples, and consecutive detection of two or more diagnostic or prognostic markers in distinct but matched samples. As used herein the term "matched samples" shall be understood to mean two or more samples derived from the same initial biological sample, or two or more biological samples isolated at the same point in time.

Accordingly, a multiplexed assay may comprise an assay that detects several antibodies and/or epitopes in the same reaction and simultaneously, or alternatively, it may detect other one or more antigens/antibodies in addition to one or more antibodies and/or epitopes. As will be apparent to the skilled artisan, if such an assay is antibody or ligand based, both of these antibodies must function under the same conditions.

#### *Diagnostic assay kits*

A further aspect of the present invention provides a kit for detecting *M. tuberculosis* infection in a biological sample. In one embodiment, the kit comprises:

- (i) one or more isolated antibodies that bind to an ovarian cancer-associated protein or an immunogenic fragment or epitope thereof; and
- (ii) means for detecting the formation of an antigen-antibody complex.

In an alternative embodiment, the kit comprises:

- (i) an isolated or recombinant ovarian cancer-associated protein or an immunogenic fragment or epitope thereof; and
- 5 (ii) means for detecting the formation of an antigen-antibody complex.

Optionally, the kit further comprises means for the detection of the binding of an antibody, fragment thereof or a ligand to an ovarian cancer-associated protein. Such means include a reporter molecule such as, for example, an enzyme (such as  
10 horseradish peroxidase or alkaline phosphatase), a substrate, a cofactor, an inhibitor, a dye, a radionucleotide, a luminescent group, a fluorescent group, biotin or a colloidal particle, such as colloidal gold or selenium. Preferably such a reporter molecule is directly linked to the antibody or ligand.

15 In yet another embodiment, a kit may additionally comprise a reference sample. Such a reference sample.

In another embodiment, a reference sample comprises a peptide that is detected by an antibody or a ligand. Preferably, the peptide is of known concentration. Such a peptide  
20 is of particular use as a standard. Accordingly various known concentrations of such a peptide may be detected using a prognostic or diagnostic assay described herein.

In yet another embodiment, a kit comprises means for protein isolation (Scopes (In:  
Protein Purification: Principles and Practice, Third Edition, Springer Verlag, 1994).

25

#### *Bioinformatics*

The ability to identify genes that are over or under expressed in ovarian cancer can additionally provide high-resolution, high-sensitivity datasets which are used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure,  
30 biosensor development, and other related areas. For example, the expression profiles are used in diagnostic or prognostic evaluation of patients with ovarian cancer. Or as another example, subcellular toxicological information are generated to better direct drug structure and activity correlation (see Anderson, *Pharmaceutical Proteomics: Targets, Mechanism, and Function*, paper presented at the IBC Proteomics conference,  
35 Coronado, CA (June 11-12, 1998)). Subcellular toxicological information can also be utilized in a biological sensor device to predict the likely toxicological effect of chemical

exposures and likely tolerable exposure thresholds (see U.S. Patent No. 5,811,231). Similar advantages accrue from datasets relevant to other biomolecules and bioactive agents (e.g., nucleic acids, saccharides, lipids, drugs, and the like).

- 5 Thus, in another embodiment, the present invention provides a database that includes at least one set of assay data. The data contained in the database is acquired, e.g., using array analysis either singly or in a library format. The database are in substantially any form in which data are maintained and transmitted, but is preferably an electronic database. The electronic database of the invention are maintained on any electronic  
10 device allowing for the storage of and access to the database, such as a personal computer, but is preferably distributed on a wide area network, such as the World Wide Web.

- The focus of the present section on databases that include peptide sequence data is for  
15 clarity of illustration only. It will be apparent to those of skill in the art that similar databases are assembled for any assay data acquired using an assay of the invention.

- The compositions and methods for identifying and/or quantitating the relative and/or absolute abundance of a variety of molecular and macromolecular species from a  
20 biological sample undergoing ovarian cancer, i.e., the identification of ovarian cancer-associated sequences described herein, provide an abundance of information, which are correlated with pathological conditions, predisposition to disease, drug testing, therapeutic monitoring, gene-disease causal linkages, identification of correlates of immunity and physiological status, among others. Although the data generated from the  
25 assays of the invention is suited for manual review and analysis, in a preferred embodiment, prior data processing using high-speed computers is utilized.

- An array of methods for indexing and retrieving biomolecular information is known in the art. For example, U.S. Patents 6,023,659 and 5,966,712 disclose a relational database  
30 system for storing biomolecular sequence information in a manner that allows sequences to be catalogued and searched according to one or more protein function hierarchies. U.S. Patent 5,953,727 discloses a relational database having sequence records containing information in a format that allows a collection of partial-length DNA sequences to be catalogued and searched according to association with one or more  
35 sequencing projects for obtaining full-length sequences from the collection of partial length sequences. U.S. Patent 5,706,498 discloses a gene database retrieval system for

- making a retrieval of a gene sequence similar to a sequence data item in a gene database based on the degree of similarity between a key sequence and a target sequence. U.S. Patent 5,538,897 discloses a method using mass spectroscopy fragmentation patterns of peptides to identify amino acid sequences in computer databases by comparison of predicted mass spectra with experimentally-derived mass spectra using a closeness-of-fit measure. U.S. Patent 5,926,818 discloses a multi-dimensional database comprising a functionality for multi-dimensional data analysis described as on-line analytical processing (OLAP), which entails the consolidation of projected and actual data according to more than one consolidation path or dimension.
- 5 U.S. Patent 5,295,261 reports a hybrid database structure in which the fields of each database record are divided into two classes, navigational and informational data, with navigational fields stored in a hierarchical topological map which are viewed as a tree structure or as the merger of two or more such tree structures.
- 10
- 15 See also Mount *et al.*, *Bioinformatics* (2001); *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids* (Durbin *et al.*, eds., 1999); *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins* (Baxevanis & Ouellette eds., 1998); Rashidi & Buehler, *Bioinformatics: Basic Applications in Biological Science and Medicine* (1999); *Introduction to Computational Molecular Biology* (Setubal *et al.*, eds
- 20 1997); *Bioinformatics: Methods and Protocols* (Misener & Krawetz, eds, 2000); *Bioinformatics: Sequence, Structure, and Databanks: A Practical Approach* (Higgins & Taylor, eds., 2000); Brown, *Bioinformatics: A Biologist's Guide to Biocomputing and the Internet* (2001); Han & Kamber, *Data Mining: Concepts and Techniques* (2000); and Waterman, *Introduction to Computational Biology: Maps, Sequences, and Genomes*
- 25 (1995).

The present invention provides a computer database comprising a computer and software for storing in computer-retrievable form assay data records cross-tabulated, e.g., with data specifying the source of the target-containing sample from which each

30 sequence specificity record was obtained.

In an exemplary embodiment, at least one of the sources of target-containing sample is from a control tissue sample known to be free of pathological disorders. In a variation, at least one of the sources is a known pathological tissue specimen, e.g., a neoplastic

35 lesion or another tissue specimen to be analyzed for prostate cancer. In another variation, the assay records cross-tabulate one or more of the following parameters for

each target species in a sample: (1) a unique identification code, which can include, e.g., a target molecular structure and/or characteristic separation coordinate (e.g., electrophoretic coordinates); (2) sample source; and (3) absolute and/or relative quantity of the target species present in the sample.

5

The invention also provides for the storage and retrieval of a collection of target data in a computer data storage apparatus, which can include magnetic disks, optical disks, magneto-optical disks, DRAM, SRAM, SGRAM, SDRAM, RDRAM, DDR RAM, magnetic bubble memory devices, and other data storage devices, including CPU registers and  
10 on-CPU data storage arrays. Typically, the target data records are stored as a bit pattern in an array of magnetic domains on a magnetizable medium or as an array of charge states or transistor gate states, such as an array of cells in a DRAM device (e.g., each cell comprised of a transistor and a charge storage area, which are on the transistor). In one embodiment, the invention provides such storage devices, and computer systems  
15 built therewith, comprising a bit pattern encoding a protein expression fingerprint record comprising unique identifiers for at least 10 target data records cross-tabulated with target source.

When the target is a peptide or nucleic acid, the invention preferably provides a method  
20 for identifying related peptide or nucleic acid sequences, comprising performing a computerised comparison between a peptide or nucleic acid sequence assay record stored in or retrieved from a computer storage device or database and at least one other sequence. The comparison can include a sequence analysis or comparison algorithm or computer program embodiment thereof (e.g., BLAST, FASTA, TFASTA, GAP, BESTFIT  
25 – see above) and/or the comparison are of the relative amount of a peptide or nucleic acid sequence in a pool of sequences determined from a polypeptide or nucleic acid sample of a specimen.

The invention also preferably provides a magnetic disk, such as an IBM-compatible  
30 (DOS, Windows, Windows95/98/2000, Windows NT, OS/2) or other format (e.g., Linux, SunOS, Solaris, AIX, SCO Unix, VMS, MV; Macintosh, etc.) floppy diskette or hard (fixed, Winchester) disk drive, comprising a bit pattern encoding data from an assay of the invention in a file format suitable for retrieval and processing in a computerized sequence analysis, comparison, or relative quantitation method.

35

The invention also provides a network, comprising a plurality of computing devices linked via a data link, such as an Ethernet cable (coax or IOBaseT), telephone line, ISDN line, wireless network, optical fiber, or other suitable signal transmission medium, whereby at least one network device (e.g., computer, disk array, etc.) comprises a pattern of  
5 magnetic domains (e.g., magnetic disk) and/or charge domains (e.g., an array of DRAM cells) composing a bit pattern encoding data acquired from an assay of the invention.

The invention also provides a method for transmitting assay data that includes generating an electronic signal on an electronic communications device, such as a  
10 modem, ISDN terminal adapter, DSL, cable modem, ATM switch, or the like, wherein the signal includes (in native or encrypted format) a bit pattern encoding data from an assay or a database comprising a plurality of assay results obtained by the method of the invention.

15 In a preferred embodiment, the invention provides a computer system for comparing a query target to a database containing an array of data structures, such as an assay result obtained by the method of the invention, and ranking database targets based on the degree of identity and gap weight to the target data. A central processor is preferably initialized to load and execute the computer program for alignment and/or comparison of  
20 the assay results. Data for a query target is entered into the central processor via an I/O device. Execution of the computer program results in the central processor retrieving the assay data from the data file, which comprises a binary description of an assay result.

The target data or record and the computer program are transferred to secondary  
25 memory, which is typically random access memory (e.g., DRAM, SRAM, SGRAM, or SDRAM). Targets are ranked according to the degree of correspondence between a selected assay characteristic (e.g., binding to a selected affinity moiety) and the same characteristic of the query target and results are output via an I/O device. For example, a central processor are a conventional computer (e.g., Intel Pentium, PowerPC, Alpha,  
30 PA-8000, SPARC, MIPS 4400, MIPS 10000, VAX, etc.); a program are a commercial or public domain molecular biology software package (e.g., UWGCG Sequence Analysis Software, Darwin); a data file are an optical or magnetic disk, a data server, a memory device (e.g., DRAM, SRAM, SGRAM, SDRAM, EPROM, bubble memory, flash memory, etc.); an I/O device are a terminal comprising a video display and a keyboard, a modem,  
35 an ISDN terminal adapter, an Ethernet port, a punched card reader, a magnetic strip reader, or other suitable I/O device.

The invention also preferably provides the use of a computer system, such as that described above, which comprises: (1) a computer; (2) a stored bit pattern encoding a collection of peptide sequence specificity records obtained by the methods of the invention, which are stored in the computer; (3) a comparison target, such as a query target; and (4) a program for alignment and comparison, typically with rank-ordering of comparison results on the basis of computed similarity values.

*Transgenic Animals Expressing Ovarian Cancer-associated proteins and "Knock-Out"*

*Animals*

The present invention also contemplates transgenic animals which are transgenic by virtue of comprising a polynucleotide of the invention, i.e. animals transformed with a cancer-associated gene of the invention. Suitable animals are generally from the phylum chordata. Chordates includes vertebrate groups such as mammals, birds, reptiles and amphibians. Particular examples of mammals include non-human primates, cats, dogs, ungulates such as cows, goats, pigs, sheep and horses and rodents such as mice, rats, gerbils and hamsters. Transgenic animals within the meaning of the present invention are non-human animals and the production of transgenic humans is specifically excluded.

Techniques for producing transgenic animals are well known in the art. A useful general textbook on this subject is Houdebine, *Transgenic animals – Generation and Use* (Harwood Academic, 1997) – an extensive review of the techniques used to generate transgenic animals from fish to mice and cows.

Advances in technologies for embryo micromanipulation now permit introduction of heterologous DNA into, for example, fertilized mammalian ova. For instance, totipotent or pluripotent stem cells are transformed by microinjection, calcium phosphate mediated precipitation, liposome fusion, retroviral infection or other means, the transformed cells are then introduced into the embryo, and the embryo then develops into a transgenic animal. In a highly preferred method, developing embryos are infected with a retrovirus containing the desired DNA, and transgenic animals produced from the infected embryo. In a most preferred method, however, the appropriate DNAs are coinjected into the pronucleus or cytoplasm of embryos, preferably at the single cell stage, and the embryos allowed to develop into mature transgenic animals. Those techniques as well known. See reviews of standard laboratory procedures for microinjection of heterologous DNAs into

mammalian fertilized ova, including Hogan *et al.*, *Manipulating the Mouse Embryo*, (Cold Spring Harbor Press 1986); Krimpenfort *et al.*, *Bio/Technology* 9:844 (1991); Palmiter *et al.*, *Cell*, 41: 343 (1985); Kraemer *et al.*, *Genetic manipulation of the Mammalian Embryo*, (Cold Spring Harbor Laboratory Press 1985); Hammer *et al.*, *Nature*, 315: 680 (1985);  
5 Wagner *et al.*, U.S. Pat. No. 5,175,385; Krimpenfort *et al.*, U.S. Pat. No. 5,175,384, the respective contents of which are incorporated herein by reference

Another method used to produce a transgenic animal involves microinjecting a nucleic acid into pro-nuclear stage eggs by standard methods. Injected eggs are then cultured  
10 before transfer into the oviducts of pseudopregnant recipients.

Transgenic animals may also be produced by nuclear transfer technology as described in Schnieke, A.E. *et al.*, 1997, *Science*, 278: 2130 and Cibelli, J.B. *et al.*, 1998, *Science*, 280: 1256. Using this method, fibroblasts from donor animals are stably transfected with  
15 a plasmid incorporating the coding sequences for a binding domain or binding partner of interest under the control of regulatory. Stable transfectants are then fused to enucleated oocytes, cultured and transferred into female recipients.

Analysis of animals which may contain transgenic sequences would typically be  
20 performed by either PCR or Southern blot analysis following standard methods.

By way of a specific example for the construction of transgenic mammals, such as cows, nucleotide constructs comprising a sequence encoding a binding domain fused to GFP are microinjected using, for example, the technique described in U.S. Pat. No. 4,873,191,  
25 into oocytes which are obtained from ovaries freshly removed from the mammal. The oocytes are aspirated from the follicles and allowed to settle before fertilization with thawed frozen sperm capacitated with heparin and prefractionated by Percoll gradient to isolate the motile fraction.

30 The fertilized oocytes are centrifuged, for example, for eight minutes at 15,000 g to visualize the pronuclei for injection and then cultured from the zygote to morula or blastocyst stage in oviduct tissue-conditioned medium. This medium is prepared by using luminal tissues scraped from oviducts and diluted in culture medium. The zygotes must be placed in the culture medium within two hours following microinjection.

35



Oestrous is then synchronized in the intended recipient mammals, such as cattle, by administering coprostanol. Oestrous is produced within two days and the embryos are transferred to the recipients 5-7 days after estrous. Successful transfer are evaluated in the offspring by Southern blot.

5

Alternatively, the desired constructs are introduced into embryonic stem cells (ES cells) and the cells cultured to ensure modification by the transgene. The modified cells are then injected into the blastula embryonic stage and the blastulas replaced into pseudopregnant hosts. The resulting offspring are chimeric with respect to the ES and host cells, and nonchimeric strains which exclusively comprise the ES progeny are obtained using conventional cross-breeding. This technique is described, for example, in WO91/10741.

10

In another embodiment, transgenic animals of the present invention are transgenic "knock-out" animals where a specific gene corresponding to a polynucleotide referred to in Tables 1-3 has been rendered non-functional by homologous recombination. The generation of "knock-out" animals is similar to the production of other transgenic animals except that the polynucleotide constructs are designed to integrate into the endogenous genes and disrupt the function of the endogenous sequences. The generation of "knock-out" animals is known in the art, including the design of suitable constructs that will recombine at the appropriate site in the genome.

15

20

In one embodiment, the heterologous sequence which it is desired to recombine into the genome of a target animal comprises a functional sequence but under the control of an inducible promoter so that expression of the gene are regulated by administration of an endogenous molecule. This are advantageous where disruption of the gene is embryonic-lethal.

25

"Knock-out" animals are used as animal models for the study of gene function.

30

#### *Therapeutic peptides*

In accordance with this embodiment, ovarian cancer-associated proteins of the present invention are administered therapeutically to patients for a time and under conditions sufficient to ameliorate the growth of a tumor in the subject or to prevent tumor recurrence.

35

It is preferred to use peptides that do not consisting solely of naturally-occurring amino acids but which have been modified, for example to reduce immunogenicity, to increase circulatory half-life in the body of the patient, to enhance bioavailability and/or to enhance efficacy and/or specificity.

5

A number of approaches have been used to modify peptides for therapeutic application. One approach is to link the peptides or proteins to a variety of polymers, such as polyethylene glycol (PEG) and polypropylene glycol (PPG) – see for example U.S. Patent Nos. 5,091,176, 5,214,131 and US 5,264,209.

10

Replacement of naturally-occurring amino acids with a variety of uncoded or modified amino acids such as D-amino acids and N-methyl amino acids may also be used to modify peptides

15 Another approach is to use bifunctional crosslinkers, such as N-succinimidyl 3-(2 pyridyldithio) propionate, succinimidyl 6-[3-(2 pyridyldithio) propionamido] hexanoate, and sulfosuccinimidyl 6-[3-(2 pyridyldithio) propionamido]hexanoate (see US Patent 5,580,853).

20 It are desirable to use derivatives of the ovarian cancer-associated proteins of the invention which are conformationally constrained. Conformational constraint refers to the stability and preferred conformation of the three-dimensional shape assumed by a peptide. Conformational constraints include local constraints, involving restricting the conformational mobility of a single residue in a peptide; regional constraints, involving  
25 restricting the conformational mobility of a group of residues, which residues may form some secondary structural unit; and global constraints, involving the entire peptide structure.

The active conformation of the peptide are stabilized by a covalent modification, such as  
30 cyclization or by incorporation of gamma-lactam or other types of bridges. For example, side chains are cyclized to the backbone so as create a L-gamma-lactam moiety on each side of the interaction site. See, generally, Hruby et al., "Applications of Synthetic Peptides," In Synthetic Peptides: A User's Guide: 259-345 (W. H. Freeman & Co. 1992). Cyclization also are achieved, for example, by formation of cystine bridges, coupling of  
35 amino and carboxy terminal groups of respective terminal amino acids, or coupling of the amino group of a Lys residue or a related homolog with a carboxy group of Asp, Glu or a

related homolog. Coupling of the .alpha-amino group of a polypeptide with the epsilon-amino group of a lysine residue, using iodoacetic anhydride, are also undertaken. See Wood and Wetzel, 1992, Int'l J. Peptide Protein Res. 39: 533-39.

- 5 Another approach described in US 5,891,418 is to include a metal-ion complexing backbone in the peptide structure. Typically, the preferred metal-peptide backbone is based on the requisite number of particular coordinating groups required by the coordination sphere of a given complexing metal ion. In general, most of the metal ions that may prove useful have a coordination number of four to six. The nature of the
- 10 coordinating groups in the peptide chain includes nitrogen atoms with amine, amide, imidazole, or guanidino functionalities; sulfur atoms of thiols or disulfides; and oxygen atoms of hydroxy, phenolic, carbonyl, or carboxyl functionalities. In addition, the peptide chain or individual amino acids are chemically altered to include a coordinating group, such as for example oxime, hydrazino, sulfhydryl, phosphate, cyano, pyridino, piperidino,
- 15 or morpholino. The peptide construct are either linear or cyclic, however a linear construct is typically preferred. One example of a small linear peptide is Gly-Gly-Gly-Gly which has four nitrogens (an  $N_4$  complexation system) in the back bone that can complex to a metal ion with a coordination number of four.
- 20 A further technique for improving the properties of therapeutic peptides is to use non-peptide peptidomimetics. A wide variety of useful techniques are used to elucidating the precise structure of a peptide. These techniques include amino acid sequencing, x-ray crystallography, mass spectroscopy, nuclear magnetic resonance spectroscopy, computer-assisted molecular modeling, peptide mapping, and combinations thereof.
- 25 Structural analysis of a peptide generally provides a large body of data which comprise the amino acid sequence of the peptide as well as the three-dimensional positioning of its atomic components. From this information, non-peptide peptidomimetics are designed that have the required chemical functionalities for therapeutic activity but are more stable, for example less susceptible to biological degradation. An example of this approach is
- 30 provided in US 5,811,512.

Techniques for chemically synthesising therapeutic peptides of the invention are described in the above references and also reviewed by Borgia and Fields, 2000, TibTech 18: 243-251 and described in detail in the references contained therein.

The ovarian cancer proteins, nucleic acids, and antibodies as described herein are used in drug screening assays to identify candidate compounds for use in treating ovarian cancer. The ovarian cancer-associated proteins, antibodies, nucleic acids, modified proteins and cells containing ovarian cancer sequences are used in drug screening assays or by evaluating the effect of drug candidates on a "gene expression profile" or expression profile of polypeptides. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent (e.g., Zlokarnik, *et al.*, 1998, *Science* 279: 84-88); Heid, 1996, *Genome Res* 6: 986-94).

In a preferred embodiment, the ovarian cancer-associated proteins, antibodies, nucleic acids, modified proteins and cells containing the native or modified ovarian cancer-associated proteins are used in screening assays. That is, the present invention provides methods for screening for compounds/agents which modulate the ovarian cancer phenotype or an identified physiological function of a ovarian cancer-associated protein. As above, this are done on an individual gene level or by evaluating the effect of drug candidates on a "gene expression profile". In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent, see Zlokarnik, *supra*.

Having identified the differentially expressed genes herein, a variety of assays are executed. In a preferred embodiment, assays are run on an individual gene or protein level. That is, having identified a particular gene as up regulated in ovarian cancer, test compounds are screened for the ability to modulate gene expression or for binding to the ovarian cancer-associated protein. "Modulation" thus includes both an increase and a decrease in gene expression. The preferred amount of modulation will depend on the original change of the gene expression in normal versus tissue undergoing ovarian cancer, with changes of at least 10%, preferably 50%, more preferably 100-300%, and in some embodiments 300-1000% or greater. Thus, if a gene exhibits a 4-fold increase in ovarian cancer tissue compared to normal tissue, a decrease of about four-fold is often desired; similarly, a 10-fold decrease in ovarian cancer tissue compared to normal tissue often provides a target value of a 10-fold increase in expression to be induced by the test compound.

The amount of gene expression are monitored using nucleic acid probes and the quantification of gene expression levels, or, alternatively, the gene product itself are monitored, e.g., through the use of antibodies to the ovarian cancer-associated protein and standard immunoassays. Proteomics and separation techniques may also allow  
5 quantification of expression.

In a preferred embodiment, gene expression or protein monitoring of a number of entities, i.e., an expression profile, is monitored simultaneously. Such profiles will typically involve a plurality of those entities described herein.

10 In this embodiment, the ovarian cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of ovarian cancer sequences in a particular cell. Alternatively, PCR are used. Thus, a series are used with dispensed primers in desired wells. A PCR reaction can then be performed and analyzed for each  
15 well.

Expression monitoring are performed to identify compounds that modify the expression of one or more ovarian cancer-associated sequences, e.g., a polynucleotide sequence set out in Tables 1-3. In a preferred embodiment, a test modulator is added to the cells  
20 prior to analysis. Moreover, screens are also provided to identify agents that modulate ovarian cancer, modulate ovarian cancer-associated proteins, bind to a ovarian cancer-associated protein, or interfere with the binding of a ovarian cancer-associated protein and an antibody or other binding partner.

25 The term "test compound" or "drug candidate" or "modulator" or grammatical equivalents as used herein describes any molecule, e.g., protein, oligopeptide, small organic molecule, polysaccharide, polynucleotide, etc., to be tested for the capacity to directly or indirectly alter the ovarian cancer phenotype or the expression of a ovarian cancer sequence, e.g., a nucleic acid or protein sequence. In preferred embodiments,  
30 modulators alter expression profiles, or expression profile nucleic acids or proteins provided herein. In one embodiment, the modulator suppresses a ovarian cancer phenotype, e.g. to a normal tissue fingerprint. In another embodiment, a modulator induced a ovarian cancer phenotype. Generally, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various  
35 concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration or below the level of detection.

Drug candidates encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 100 and less than about 2,500 daltons. Preferred small molecules are less than 2000, or less than 1500 or less than 1000 or less than 500 Daltons. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof. Particularly preferred are peptides.

In one aspect, a modulator will neutralize the effect of a ovarian cancer-associated protein. By "neutralize" is meant that activity of a protein is inhibited or blocked and the consequent effect on the cell.

In certain embodiments, combinatorial libraries of potential modulators will be screened for an ability to bind to a ovarian cancer polypeptide or to modulate activity. Conventionally, new chemical entities with useful properties are generated by identifying a chemical compound (called a "lead compound") with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

In one preferred embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate compounds). Such "combinatorial chemical libraries" are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical

"building blocks" such as reagents. For example, a linear combinatorial chemical library, such as a polypeptide (e.g., mutein) library, is formed by combining a set of chemical building blocks called amino acids in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds are synthesized through such combinatorial mixing of chemical building blocks (Gallop *et al.*, 1994, *J. Med. Chem.* 37(9):1233-1251).

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries, peptoids, encoded peptides, random bio-oligomers, nonpeptidal peptidomimetics, analogous organic syntheses of small compound libraries, nucleic acid libraries, peptide nucleic acid libraries, antibody libraries, carbohydrate libraries and small organic molecule libraries.

The assays to identify modulators are amenable to high throughput screening. Preferred assays thus detect enhancement or inhibition of ovarian cancer gene transcription, inhibition or enhancement of polypeptide expression, and inhibition or enhancement of polypeptide activity.

High throughput assays for the presence, absence, quantification, or other properties of particular nucleic acids or protein products are well known to those of skill in the art. Similarly, binding assays and reporter gene assays are similarly well known. Thus, e.g., U.S. Patent No. 5,559,410 discloses high throughput screening methods for proteins, U.S. Patent No. 5,585,639 discloses high throughput screening methods for nucleic acid binding (i.e., in arrays), while U.S. Patent Nos. 5,576,220 and 5,541,061 disclose high throughput methods of screening for ligand/antibody binding.

In addition, high throughput screening systems are commercially available (see, e.g., Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, etc.). These systems typically automate entire procedures, including all sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detectors) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization. The manufacturers of such systems provide detailed protocols for various high throughput systems. Thus, e.g.,

Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like.

5 In one embodiment, modulators are proteins, often naturally occurring proteins or fragments of naturally occurring proteins. Thus, e.g., cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, are used. In this way libraries of proteins are made for screening in the methods of the invention. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred.  
10 Particularly useful test compound will be directed to the class of proteins to which the target belongs, e.g., substrates for enzymes or ligands and receptors.

In a preferred embodiment, modulators are peptides of from about 5 to about 30 amino acids, with from about 5 to about 20 amino acids being preferred, and from about 7 to  
15 about 15 being particularly preferred. The peptides are digests of naturally occurring proteins as is outlined above, random peptides, or "biased" random peptides. By "randomized" or grammatical equivalents herein is meant that each nucleic acid and peptide consists of essentially random nucleotides and amino acids, respectively. Since generally these random peptides (or nucleic acids, discussed below) are chemically  
20 synthesized, they may incorporate any nucleotide or amino acid at any position. The synthetic process are designed to generate randomized proteins or nucleic acids, to allow the formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized candidate bioactive proteinaceous agents.

25 In one embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a limited number of possibilities. For example, in a preferred embodiment, the nucleotides or  
30 amino acid residues are randomized within a defined class, e.g., of hydrophobic amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of nucleic acid binding domains, the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc., or to purines, etc.

35



Modulators of ovarian cancer can also be nucleic acids, as defined below. As described above generally for proteins, nucleic acid modulating agents are naturally occurring nucleic acids, random nucleic acids, or "biased" random nucleic acids. For example, digests of procaryotic or eucaryotic genomes are used as is outlined above for proteins.

5

In certain embodiments, the activity of a ovarian cancer-associated protein is down-regulated, or entirely inhibited, by the use of antisense polynucleotide, *i.e.*, a nucleic acid complementary to, and which can preferably hybridize specifically to, a coding mRNA nucleic acid sequence, *e.g.*, a ovarian cancer-associated protein mRNA, or a subsequence thereof. Binding of the antisense polynucleotide to the mRNA reduces the translation and/or stability of the mRNA.

10

In the context of this invention, antisense polynucleotides can comprise naturally-occurring nucleotides, or synthetic species formed from naturally-occurring subunits or their close homologs. Antisense polynucleotides may also have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species which are known for use in the art. Analogs are comprehended by this invention so long as they function effectively to hybridize with the ovarian cancer-associated protein mRNA. See, *e.g.*, Isis Pharmaceuticals, Carlsbad, CA; Sequitor, Inc., Natick, MA.

20

Such antisense polynucleotides can readily be synthesized using recombinant means, or are synthesized *in vitro*. Equipment for such synthesis is sold by several vendors, including Applied Biosystems. The preparation of other oligonucleotides such as phosphorothioates and alkylated derivatives is also well known to those of skill in the art.

25

Antisense molecules as used herein include antisense or sense oligonucleotides. Sense oligonucleotides can, *e.g.*, be employed to block transcription by binding to the anti-sense strand. The antisense and sense oligonucleotide comprise a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target mRNA (sense) or DNA (antisense) sequences for ovarian cancer molecules. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment generally at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, *e.g.*, Stein & Cohen (*Cancer Res.* 48:2659 (1988) and van der Krol *et al.* (*BioTechniques* 6:958 (1988)).

30

35

- In addition to antisense polynucleotides, ribozymes are used to target and inhibit transcription of ovarian cancer-associated nucleotide sequences. A ribozyme is an RNA molecule that catalytically cleaves other RNA molecules. Different kinds of ribozymes have been described, including group I ribozymes, hammerhead ribozymes, hairpin ribozymes, RNase P, and axhead ribozymes (see, e.g., Castanotto *et al.*, *Adv. in Pharmacology* 25: 289-317 (1994) for a general review of the properties of different 5 ribozymes).
- 10 Methods of preparing ribozymes are well known to those of skill in the art (see, e.g., WO 94/26877; Ojwang *et al.*, *Proc. Natl. Acad. Sci. USA* 90:6340-6344 (1993); Yamada *et al.*, *Human Gene Therapy* 1:39-45 (1994); Leavitt *et al.*, *Proc. Natl. Acad. Sci. USA* 92:699- 703 (1995); Leavitt *et al.*, *Human Gene Therapy* 5:1151-120 (1994); and Yamada *et al.*, *Virology* 205: 121-126 (1994)).
- 15 Polynucleotide modulators of ovarian cancer are introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that
- 20 bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell. Alternatively, a polynucleotide modulator of ovarian cancer are introduced into a cell containing the target nucleic acid
- 25 sequence, e.g., by formation of an polynucleotide-lipid complex, as described in WO 90/10448. It is understood that the use of antisense molecules or knock out and knock in models may also be used in screening assays as discussed above, in addition to methods of treatment.
- 30 As noted above, gene expression monitoring is conveniently used to test candidate modulators (e.g., protein, nucleic acid or small molecule). After the candidate agent has been added and the cells allowed to incubate for some period of time, the sample containing a target sequence to be analyzed is added to the biochip. If required, the target sequence is prepared using known techniques. For example, the sample are
- 35 treated to lyse the cells, using known lysis buffers, electroporation, etc., with purification and/or amplification such as PCR performed as appropriate. For example, an *in vitro*

transcription with labels covalently attached to the nucleotides is performed. Generally, the nucleic acids are labeled with biotin-FITC or PE, or with cy3 or cy5.

5 In a preferred embodiment, the target sequence is labeled with, e.g., a fluorescent, a chemiluminescent, a chemical, or a radioactive signal, to provide a means of detecting the target sequence's specific binding to a probe. The label also are an enzyme, such as, alkaline phosphatase or horseradish peroxidase, which when provided with an appropriate substrate produces a product that are detected. Alternatively, the label are a labeled compound or small molecule, such as an enzyme inhibitor, that binds but is not  
10 catalyzed or altered by the enzyme. The label also are a moiety or compound, such as, an epitope tag or biotin which specifically binds to streptavidin. For the example of biotin, the streptavidin is labeled as described above, thereby, providing a detectable signal for the bound target sequence. Unbound labeled streptavidin is typically removed prior to analysis.

15 As will be appreciated by those in the art, these assays are direct hybridization assays or can comprise "sandwich assays", which include the use of multiple probes, as is generally outlined in U.S. Patent Nos. 5,681,702, 5,597,909, 5,545,730, 5,594,117, 5,591,584, 5,571,670, 5,580,731, 5,571,670, 5,591,584, 5,624,802, 5,635,352,  
20 5,594,118, 5,359,100, 5,124,246 and 5,681,697, all of which are hereby incorporated by reference. In this embodiment, in general, the target nucleic acid is prepared as outlined above, and then added to the biochip comprising a plurality of nucleic acid probes, under conditions that allow the formation of a hybridization complex.

25 A variety of hybridization conditions are used in the present invention, including high, moderate and low stringency conditions as outlined above. The assays are generally run under stringency conditions which allows formation of the label probe hybridization complex only in the presence of target. Stringency are controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature,  
30 formamide concentration, salt concentration, chaotropic salt concentration pH, organic solvent concentration, etc.

These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Patent No. 5,681,697. Thus it are desirable to perform certain steps at  
35 higher stringency conditions to reduce non-specific binding.

The reactions outlined herein are accomplished in a variety of ways. Components of the reaction are added simultaneously, or sequentially, in different orders, with preferred embodiments outlined below. In addition, the reaction may include a variety of other reagents. These include salts, buffers, neutral proteins, e.g. albumin, detergents, *etc.* which are used to facilitate optimal hybridization and detection, and/or reduce non-specific or background interactions. Reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, *etc.*, may also be used as appropriate, depending on the sample preparation methods and purity of the target.

The assay data are analyzed to determine the expression levels, and changes in expression levels as between states, of individual genes, forming a gene expression profile.

Screens are performed to identify modulators of the ovarian cancer phenotype. In one embodiment, screening is performed to identify modulators that can induce or suppress a particular expression profile, thus preferably generating the associated phenotype. In another embodiment, e.g., for diagnostic applications, having identified differentially expressed genes important in a particular state, screens are performed to identify modulators that alter expression of individual genes. In an another embodiment, screening is performed to identify modulators that alter a biological function of the expression product of a differentially expressed gene. Again, having identified the importance of a gene in a particular state, screens are performed to identify agents that bind and/or modulate the biological activity of the gene product.

In addition screens are done for genes that are induced in response to a candidate agent. After identifying a modulator based upon its ability to suppress a ovarian cancer expression pattern leading to a normal expression pattern, or to modulate a single ovarian cancer gene expression profile so as to mimic the expression of the gene from normal tissue, a screen as described above are performed to identify genes that are specifically modulated in response to the agent. Comparing expression profiles between normal tissue and agent treated ovarian cancer tissue reveals genes that are not expressed in normal tissue or ovarian cancer tissue, but are expressed in agent treated tissue. These agent-specific sequences are identified and used by methods described herein for ovarian cancer genes or proteins. In particular these sequences and the proteins they encode find use in marking or identifying agent treated cells. In addition,

antibodies are raised against the agent induced proteins and used to target novel therapeutics to the treated ovarian cancer tissue sample.

Thus, in one embodiment, a test compound is administered to a population of ovarian cancer cells, that have an associated ovarian cancer expression profile. By "administration" or "contacting" herein is meant that the candidate agent is added to the cells in such a manner as to allow the agent to act upon the cell, whether by uptake and intracellular action, or by action at the cell surface. In some embodiments, nucleic acid encoding a proteinaceous candidate agent (i.e., a peptide) are put into a viral construct such as an adenoviral or retroviral construct, and added to the cell, such that expression of the peptide agent is accomplished. Regulatable gene administration systems can also be used.

Once the test compound has been administered to the cells, the cells are washed if desired and are allowed to incubate under preferably physiological conditions for some period of time. The cells are then harvested and a new gene expression profile is generated, as outlined herein.

Thus, e.g., ovarian cancer tissue are screened for agents that modulate, e.g., induce or suppress the ovarian cancer phenotype. A change in at least one gene, preferably many, of the expression profile indicates that the agent has an effect on ovarian cancer activity. By defining such a signature for the ovarian cancer phenotype, screens for new drugs that alter the phenotype are devised. With this approach, the drug target need not be known and need not be represented in the original expression screening platform, nor does the level of transcript for the target protein need to change.

In a preferred embodiment, as outlined above, screens are done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of either the expression of the gene or the gene product itself are done. The gene products of differentially expressed genes are sometimes referred to herein as "ovarian cancer-associated proteins" or a "ovarian cancer modulatory protein". The ovarian cancer modulatory protein are a fragment, or alternatively, be the full length protein to the fragment encoded by the nucleic acids referred to in Tables 1-3. Preferably, the ovarian cancer modulatory protein is a fragment. In a preferred embodiment, the ovarian cancer amino acid sequence which is used to determine sequence identity or similarity is encoded by a

nucleic acid referred to in Tables 1-3. In another embodiment, the sequences are naturally occurring allelic variants of a protein encoded by a nucleic acid referred to in Tables 1-3. In another embodiment, the sequences are sequence variants as further described herein.

5

Preferably, the ovarian cancer modulatory protein is a fragment of approximately 14 to 24 amino acids long. More preferably the fragment is a soluble fragment. Preferably, the fragment includes a non-transmembrane region. In a preferred embodiment, the fragment has an N-terminal Cys to aid in solubility. In one embodiment, the C-terminus of the fragment is kept as a free acid and the N-terminus is a free amine to aid in coupling, i.e., to cysteine.

10

In one embodiment the ovarian cancer-associated proteins are conjugated to an immunogenic agent as discussed herein. In one embodiment the ovarian cancer-associated protein is conjugated to BSA.

15

Measurements of ovarian cancer polypeptide activity, or of ovarian cancer or the ovarian cancer phenotype are performed using a variety of assays. For example, the effects of the test compounds upon the function of the ovarian cancer polypeptides are measured by examining parameters described above. A suitable physiological change that affects activity are used to assess the influence of a test compound on the polypeptides of this invention. When the functional consequences are determined using intact cells or animals, one can also measure a variety of effects such as, in the case of ovarian cancer associated with tumours, tumour growth, tumour metastasis, neovascularization, hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second messengers such as cGMP. In tire assays of the invention, mammalian ovarian cancer polypeptide is typically used, e.g., mouse, preferably human.

20

25

30

Assays to identify compounds with modulating activity are performed *in vitro*. For example, a ovarian cancer polypeptide is first contacted with a potential modulator and incubated for a suitable amount of time, e.g., from 0.5 to 48 hours. In one embodiment, the ovarian cancer polypeptide levels are determined *in vitro* by measuring the level of protein or mRNA. The level of protein is measured using immunoassays such as western blotting, ELISA and the like with an antibody that selectively binds to the ovarian cancer

35

polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using PCR, LCR, or hybridization assays; e.g., northern hybridization, RNase protection, dot blotting, are preferred. The level of protein or mRNA is detected using directly or indirectly labeled detection agents, e.g., fluorescently or radioactively labeled nucleic acids, radioactively or enzymatically labeled antibodies, and the like, as described herein.

Alternatively, a reporter gene system are devised using the ovarian cancer-associated protein promoter operably linked to a reporter gene such as luciferase, green fluorescent protein, CAT, or (beta-gal. The reporter construct is typically transfected into a cell. After treatment with a potential modulator, the amount of reporter gene transcription, translation, or activity is measured according to standard techniques known to those of skill in the art.

In a preferred embodiment, as outlined above, screens are done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of the expression of the gene or the gene product itself are done. The gene products of differentially expressed genes are sometimes referred to herein as "ovarian cancer-associated proteins." The ovarian cancer-associated protein are a fragment, or alternatively, be the full length protein to a fragment shown herein.

In one embodiment, screening for modulators of expression of specific genes is performed. Typically, the expression of only one or a few genes are evaluated. In another embodiment, screens are designed to first find compounds that bind to differentially expressed proteins. These compounds are then evaluated for the ability to modulate differentially expressed activity. Moreover, once initial candidate compounds are identified, variants are further screened to better evaluate structure activity relationships.

In a preferred embodiment, binding assays are done. In general, purified or isolated gene product is used; that is, the gene products of one or more differentially expressed nucleic acids are made. For example, antibodies are generated to the protein gene products, and standard immunoassays are run to determine the amount of protein present. Alternatively, cells comprising the ovarian cancer-associated proteins are used in the assays.

Thus, in a preferred embodiment, the methods comprise combining a ovarian cancer-associated protein and a candidate compound, and determining the binding of the compound to the ovarian cancer-associated protein. Preferred embodiments utilize the human ovarian cancer-associated protein, although other mammalian proteins may also  
5 be used, e.g. for the development of animal models of human disease. In some embodiments, as outlined herein, variant or derivative ovarian cancer-associated proteins are used.

Generally, in a preferred embodiment of the methods herein, the ovarian cancer-associated protein or the candidate agent is non-diffusably bound to an insoluble support  
10 having isolated sample receiving areas (e.g. a microtiter plate, an array, etc.). The insoluble supports are made of any composition to which the compositions are bound, is readily separated from soluble material, and is otherwise compatible with the overall method of screening. The surface of such supports are solid or porous and of any  
15 convenient shape. Examples of suitable insoluble supports include microtiter plates, arrays, membranes and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharides, nylon or nitrocellulose, teflon™, etc. microtitre plates and arrays are especially convenient because a large number of assays are carried out simultaneously, using small amounts of reagents and samples. The particular manner of  
20 binding of the composition is not crucial so long as it is compatible with the reagents and overall methods of the invention, maintains the activity of the composition and is nondiffusable. Preferred methods of binding include the use of antibodies (which do not sterically block either the ligand binding site or activation sequence when the protein is bound to the support), direct binding to "sticky" or ionic supports, chemical crosslinking,  
25 the synthesis of the protein or agent on the surface, etc. Following binding of the protein or agent, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein or other innocuous protein or other moiety.

30 In a preferred embodiment, the ovarian cancer-associated protein is bound to the support, and a test compound is added to the assay. Alternatively, the candidate agent is bound to the support and the ovarian cancer-associated protein is added. Novel binding agents include specific antibodies, non-natural binding agents identified in screens of chemical libraries, peptide analogs, etc. Of particular interest are screening assays for  
35 agents that have a low toxicity for human cells. A wide variety of assays are used for this purpose, including labeled in vitro protein-protein binding assays, electrophoretic mobility



shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.) and the like.

5 The determination of the binding of the test modulating compound to the ovarian cancer-associated protein are done in a number of ways. In a preferred embodiment, the compound is labeled, and binding determined directly, e.g., by attaching all or a portion of the ovarian cancer-associated protein to a solid support, adding a labeled candidate agent (e.g., a fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps are utilized  
10 as appropriate.

In some embodiments, only one of the components is labeled, e.g., the proteins (or proteinaceous candidate compounds) are labeled. Alternatively, more than one component are labeled with different labels, e.g.,  $^{125}\text{I}$  for the proteins and a fluorophor for  
15 the compound. Proximity reagents, e.g., quenching or energy transfer reagents are also useful.

In one embodiment, the binding of the test compound is determined by competitive binding assay. The competitor is a binding moiety known to bind to the target molecule  
20 (i.e., a ovarian cancer-associated protein), such as an antibody, peptide, binding partner, ligand, etc. Under certain circumstances, there are competitive binding between the compound and the binding moiety, with the binding moiety displacing the compound. In one embodiment, the test compound is labeled. Either the compound, or the competitor, or both, is added first to the protein for a time sufficient to allow binding, if present.  
25 Incubations are performed at a temperature which facilitates optimal activity, typically between 4 and 40°C. Incubation periods are typically optimized, e.g., to facilitate rapid high throughput screening. Typically between 0.1 and 1 hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

30 In a preferred embodiment, the competitor is added first, followed by the test compound. Displacement of the competitor is an indication that the test compound is binding to the ovarian cancer-associated protein and thus is capable of binding to, and potentially modulating, the activity of the ovarian cancer-associated protein. In this embodiment,  
35 either component are labeled. Thus, e.g., if the competitor is labeled, the presence of

label in the wash solution indicates displacement by the agent. Alternatively, if the test compound is labeled, the presence of the label on the support indicates displacement.

5 In an alternative preferred embodiment, the test compound is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor may indicate that the test compound is bound to the ovarian cancer-associated protein with a higher affinity. Thus, if the test compound is labeled, the presence of the label on the support, coupled with a lack of competitor binding, may indicate that the test compound is capable of binding to the ovarian cancer-associated protein.

10

In a preferred embodiment, the methods comprise differential screening to identify agents that are capable of modulating the activity of the ovarian cancer-associated proteins. In this embodiment, the methods comprise combining a ovarian cancer-associated protein and a competitor in a first sample. A second sample comprises a test  
15 compound, a ovarian cancer-associated protein, and a competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the ovarian cancer-associated protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the  
20 agent is capable of binding to the ovarian cancer-associated protein.

Alternatively, differential screening is used to identify drug candidates that bind to the native ovarian cancer-associated protein, but cannot bind to modified ovarian cancer-associated proteins. The structure of the ovarian cancer-associated protein are modeled,  
25 and used in rational drug design to synthesize agents that interact with that site. Drug candidates that affect the activity of a ovarian cancer-associated protein are also identified by screening drugs for the ability to either enhance or reduce the activity of the protein.

30 Positive controls and negative controls are used in the assays. Preferably control and test samples are performed in at least triplicate to obtain statistically significant results. Incubation of all samples is for a time sufficient for the binding of the agent to the protein. Following incubation, samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a  
35 radiolabel is employed, the samples are counted in a scintillation counter to determine the amount of bound compound.

A variety of other reagents are included in the screening assays. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc. which are used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions.

- 5 Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., are used. The mixture of components are added in an order that provides for the requisite binding.

- 10 In a preferred embodiment, the invention provides methods for screening for a compound capable of modulating the activity of a ovarian cancer-associated protein. The methods comprise adding a test compound, as defined above, to a cell comprising ovarian cancer-associated proteins. Preferred cell types include almost any cell. The cells contain a recombinant nucleic acid that encodes a ovarian cancer-associated protein. In a preferred embodiment, a library of candidate agents are tested on a plurality of cells.

- 15 In one aspect, the assays are evaluated in the presence or absence or previous or subsequent exposure of physiological signals, e.g. hormones, antibodies, peptides, antigens, cytokines, growth factors, action potentials, pharmacological agents including chemotherapeutics, radiation, carcinogenics, or other cells (i.e. cell-cell contacts). In another example, the determinations are determined at different stages of the cell cycle process.

- 20 In this way, compounds that modulate ovarian cancer agents are identified. Compounds with pharmacological activity are able to enhance or interfere with the activity of the ovarian cancer-associated protein. Once identified, similar structures are evaluated to identify critical structural feature of the compound.

- 30 In one embodiment, a method of inhibiting ovarian cancer cell division is provided. The method comprises administration of a ovarian cancer inhibitor. In another embodiment, a method of inhibiting ovarian cancer is provided. The method comprises administration of a ovarian cancer inhibitor. In a further embodiment, methods of treating cells or individuals with ovarian cancer are provided. The method comprises administration of a ovarian cancer inhibitor.

- 35 In one embodiment, a ovarian cancer inhibitor is an antibody as discussed above. In another embodiment, the ovarian cancer inhibitor is an antisense molecule.

A variety of cell growth, proliferation, and metastasis assays are known to those of skill in the art, as described below.

5    *Soft agar growth or colony formation in suspension*

Normal cells require a solid substrate to attach and grow. When the cells are transformed, they lose this phenotype and grow detached from the substrate. For example, transformed cells can grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft agar. The transformed cells, when  
10    transfected with tumour suppressor genes, regenerate normal phenotype and require a solid substrate to attach and grow. Soft agar growth or colony formation in suspension assays are used to identify modulators of ovarian cancer sequences, which when expressed in host cells, inhibit abnormal cellular proliferation and transformation. A therapeutic compound would reduce or eliminate the host cells' ability to grow in stirred  
15    suspension culture or suspended in semisolid media, such as semi-solid or soft.

Techniques for soft agar growth or colony formation in suspension assays are described in Freshney, *Culture of Animal Cells a Manual of Basic Technique* (3rd ed., 1994), herein incorporated by reference. See also, the methods section of Garkavtsev *et al.* (1996),  
20    supra, herein incorporated by reference.

*Contact inhibition and density limitation of growth*

Normal cells typically grow in a flat and organized pattern in a petri dish until they touch other cells. When the cells touch one another, they are contact inhibited and stop  
25    growing. When cells are transformed, however, the cells are not contact inhibited and continue to grow to high densities in disorganized foci. Thus, the transformed cells grow to a higher saturation density than normal cells. This are detected morphologically by the formation of a disoriented monolayer of cells or rounded cells in foci within the regular pattern of normal surrounding cells. Alternatively, labeling index with (<sup>3</sup>H)-thymidine at  
30    saturation density are used to measure density limitation of growth. See Freshney (1994), supra. The transformed cells, when transfected with tumour suppressor genes, regenerate a normal phenotype and become contact inhibited and would grow to a lower density.

35    In this assay, labeling index with (<sup>3</sup>H)-thymidine at saturation density is a preferred method of measuring density limitation of growth. Transformed host cells are transfected

with a ovarian cancer-associated sequence and are grown for 24 hours at saturation density in non-limiting medium conditions. The percentage of cells labeling with (<sup>3</sup>H)-thymidine is determined autoradiographically. See, Freshney (1994), *supra*.

5     *Growth factor or serum dependence*

Transformed cells have a lower serum dependence than their normal counterparts (see, e.g., Temin, J. *Natl. Cancer Insti.* 37:167-175 (1966); Eagle *et al.*, *J. Exp. Med.* 131:836-879 (1970)); Freshney, *supra*. This is in part due to release of various growth factors by the transformed cells. Growth factor or serum dependence of transformed host  
10     cells are compared with that of control. *Tumor specific markers levels* Tumor cells release an increased amount of certain factors (hereinafter "tumour specific markers") than their normal counterparts. For example, plasminogen activator (PA) is released from human glioma at a higher level than from normal brain cells (see, e.g., Gullino, *Angiogenesis, tumour vascularization, and potential interference with tumour growth.* in  
15     *Biological Responses in Cancer*, pp. 178-184 (Mihich (ed.) 1985)). Similarly, Tumor angiogenesis factor (TAF) is released at a higher level in tumour cells than their normal counterparts. See, e.g., Folkman, *Angiogenesis and Cancer*, *Sem Cancer Biol.* (1992)). Various techniques which measure the release of these factors are described in Freshney (1994), *supra*. Also, see, Unkless *et al.*, *J. Biol. Chem.* 249:4295-4305 (1974);  
20     Strickland & Beers, *J. Biol. Chem.* 251:5694-5702 (1976); Whur *et al.*, *Br. J. Cancer* 42:305 312 (1980); Gullino, *Angiogenesis, tumour vascularization, and potential interference with tumour growth.* in *Biological Responses in Cancer*, pp. 178-184 (Mihich (ed.) 1985); Freshney *Anticancer Res.* 5:111-130 (1985).

25     *Invasiveness into Matrigel*

The degree of invasiveness into Matrigel-or some other extracellular matrix constituent are used as an assay to identify compounds that modulate ovarian cancer-associated sequences. Tumor cells exhibit a good correlation between malignancy and invasiveness of cells into Matrigel or some other extracellular matrix constituent. In this assay,  
30     tumourigenic cells are typically used as host cells. Expression of a tumour suppressor gene in these host cells would decrease invasiveness of the host cells.

Techniques described in Freshney (1994), *supra*, are used. Briefly, the level of invasion of host cells are measured by using filters coated with Matrigel or some other  
35     extracellular matrix constituent. Penetration into the gel, or through to the distal side of the filter, is rated as invasiveness, and rated histologically by number of cells and

distance moved, or by prelabeling the cells with  $^{125}\text{I}$  and counting the radioactivity on the distal side of the filter or bottom of the dish. See, e.g., Freshney (1984), *supra*.

#### *Tumor growth in vivo*

- 5 Effects of ovarian cancer-associated sequences on cell growth are tested in transgenic or immune-suppressed mice. Knock-out transgenic mice are made, in which the ovarian cancer gene is disrupted or in which a ovarian cancer gene is inserted. Knock- out transgenic mice are made by insertion of a marker gene or other heterologous gene into the endogenous ovarian cancer gene site in the mouse genome via homologous recombination. Such mice can also be made by substituting the endogenous ovarian cancer gene with a mutated version of the ovarian cancer gene, or by mutating the endogenous ovarian cancer gene, e.g., by exposure to carcinogens.

- 15 A DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells partially derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic lesion (see, e.g., Capecchi *et al.*, *Science* 244:1288 (1989)). Chimeric targeted mice are derived according to Hogan *et al.*, *Manipulating the Mouse Embryo: A Laboratory Manual*, Cold Spring Harbor Laboratory (1988) and *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, ed., IRL Press, Washington, D.C., (1987).

- 25 Alternatively, various immune-suppressed or immune-deficient host animals are used. For example, genetically athymic "nude" mouse (see, e.g., Giovanella *et al.*, *J. Natl. Cancer Inst.* 52:921 (1974)), a SCID mouse, a thymectomized mouse, or an irradiated mouse (see, e.g., Bradley *et al.*, *Br. J. Cancer* 38:263 (1978); Selby *et al.*, *Br. J. Cancer* 41:52 (1980)) are used as a host. Transplantable tumour cells (typically about  $10^8$  cells) injected into isogenic hosts will produce invasive tumours in a high proportions of cases, while normal cells of similar origin will not. In hosts which developed invasive tumours, cells expressing a ovarian cancer-associated sequences are injected subcutaneously. After a suitable length of time, preferably 4 to 8 weeks, tumour growth is measured (e.g. by volume or by its two largest dimensions) and compared to the control. Tumours that have a statistically significant reduction (using, e.g. Student's T test) are said to have inhibited growth.

*Administration*

therapeutic reagents of the invention are administered to patients, therapeutically. Typically, such proteins/polynucleotides and substances may preferably be combined  
5 with various components to produce compositions of the invention. Preferably the compositions are combined with a pharmaceutically acceptable carrier or diluent to produce a pharmaceutical composition (which are for human or animal use). Suitable carriers and diluents include isotonic saline solutions, for example phosphate-buffered saline. The composition of the invention are administered by direct injection. The  
10 composition are formulated for parenteral, intramuscular, intravenous, subcutaneous, intraocular, oral, vaginal or transdermal administration. Typically, each protein are administered at a dose of from 0.01 to 30 mg/kg body weight, preferably from 0.1 to 10 mg/kg, more preferably from 0.1 to 1 mg/kg body weight.

15 Polynucleotides/vectors encoding polypeptide components for use in modulating the activity of the ovarian cancer-associated proteins/polynucleotides are administered directly as a naked nucleic acid construct. When the polynucleotides/vectors are administered as a naked nucleic acid, the amount of nucleic acid administered may typically be in the range of from 1 µg to 10 mg, preferably from 100 µg to 1 mg.

20 Uptake of naked nucleic acid constructs by mammalian cells is enhanced by several known transfection techniques for example those including the use of transfection agents. Example of these agents include cationic agents (for example calcium phosphate and DEAE-dextran) and lipofectants (for example lipofectam<sup>TM</sup> and transfectam<sup>TM</sup>).  
25 Typically, nucleic acid constructs are mixed with the transfection agent to produce a composition.

Preferably the polynucleotide or vector of the invention is combined with a pharmaceutically acceptable carrier or diluent to produce a pharmaceutical composition.  
30 Suitable carriers and diluents include isotonic saline solutions, for example phosphate-buffered saline. The composition are formulated for parenteral, intramuscular, intravenous, subcutaneous, oral, intraocular or transdermal administration.

The pharmaceutical compositions are administered in a range of unit dosage forms  
35 depending on the method of administration. For example, unit dosage forms suitable for oral administration include, powder, tablets, pills, capsules and lozenges. Orally

administered dosage forms will typically be formulated to protect the active ingredient from digestion and may therefore be complexed with appropriate carrier molecules and/or packaged in an appropriately resistant carrier. Suitable carrier molecules and packaging materials/barrier materials are known in the art.

5

The compositions of the invention are administered for therapeutic or prophylactic treatments. In therapeutic applications, compositions are administered to a patient suffering from a disease (e.g. ovarian cancer) in an amount sufficient to cure or at least partially ameliorate the disease and its complications. An amount adequate to  
10 accomplish this is defined as a "therapeutically effective dose". An amount of the composition that is capable of preventing or slowing the development of cancer in a patient is referred to as a "prophylactically effective dose".

15

The routes of administration and dosages described are intended only as a guide since a skilled practitioner will be able to determine readily the optimum route of administration and dosage for any particular patient and condition.

20

The present invention is further described with reference to the accompanying drawings and the following non-limiting examples.



## EXAMPLE 1

*Gene expression profiling to identify differentially-expressed  
genes in ovarian cancer*

1. *Tissue Bank and Database*

5 Tissue was collected from patients undergoing treatment at the GCC, we have established an Ovarian Cancer Tissue Bank and Clinical Database that currently holds data on over 400 cases treated at the GCC between 1986 and 2002. Tissue (currently 149 fresh/frozen and 292 archival fixed paraffin-embedded samples) was acquired from patients undergoing cytoreductive surgery and does not interfere with the collection of tissue for the normal  
10 processing of diagnostic specimens. Patient consent, included in all our studies, was collected prior to surgery. Tissue specimens and their associated pathology reports were coded in order to maintain patient confidentiality. Uncoded data was electronically and/or physically locked with restricted access by appropriate senior investigators only. Clinical (diagnosis, treatment, residual disease) and pathological data (tumour grade, stage) were  
15 collected and updated (disease recurrence, patient survival) at regular intervals. This study has ethical approval from the South Eastern Sydney Area Health Service Research Ethics Committee, Australia. Clinical data and tissue collection are ongoing.

2. *Genetic profiling of ovarian cancers*

20 In order to identify those genes differentially regulated in epithelial ovarian cancer 51 ovarian cancer tumor samples were manually dissected from biological samples derived from subjects undergoing cytoreductive surgery. These samples comprised 8 endometrioid tumors, 4 mucinous tumors and 31 serous epithelial ovarian tumors, 12 corresponding omental deposits and 8 borderline (low-malignant potential) tumors.

25 RNA was isolated from the tumor samples in addition to 4 normal ovary samples using Trizol reagent (Life Technologies, Rockville, MD, USA) essentially according to manufacturer's instructions. RNA was then reverse transcribed using an oligo(dT) anchored oligonucleotide that additionally comprised a T7 promoter sequence. Isolated  
30 cDNA was then transcribed *in vitro* using the T7 MEGAscript kit (Ambion, Austin, TX, USA) according to manufacturer's instructions. Transcription was performed with biotinylated nucleotides (Bio-11-CTP and Bio-16-UTP) to enable detection of the transcribed cRNA.

35 Levels of gene expression in the cancer samples was then determined by analysing the transcribed cDNA samples using customized Affymetrix GeneChip® microarrays that

comprise 59,618 oligonucleotide probe sets. These probe sets facilitate analysis of 46,000 gene clusters, representing over 90% of the predicted expressed human genome.

5 Data were normalized, and changes in gene expression detected using a ranked penalized t-statistic with p-values adjusted for multiple testing using the Holm procedure. Analysis was performed using the LIMMA package (available from Bioconductor, Biostatistics Unit of the Dana Farber Cancer Institute at the Harvard Medical School/Harvard School of Public Health).

10 Gene expression in 186 samples representing 52 different tissues of the body was also determined using the previously described methods to facilitate the identification of changes in gene expression that are specific for ovarian cancer.

15 Using this method 284 up-regulated transcripts and 186 down-regulated transcripts were identified.

In order to determine the efficacy of such a method of analysis for determining gene expression changes associated with ovarian cancer, those genes identified were compared to results of published expression profile studies. Using this method, 71  
20 genes were identified in the present study that had been previously identified, including, for example, genes known to be over-expressed in ovarian cancer, such as, for example MUC1 and E-cadherin.

The ovarian cancer-associated genes and proteins set forth in Table 1 include  
25 sequences that are up-regulated or down-regulated in ovarian cancer subjects, including subjects suffering specifically from serous, endometrioid, mucinous or clear cell ovarian cancer, or non-invasive (borderline) ovarian cancers of any phenotype, and subjects that suffered from recurrences of ovarian cancer in the medium term, or died within the medium term.

30 Data presented in Table 2 indicate those genes that are expressed at significantly higher levels or significantly reduced levels in patients suffering from serous cancer relative to the level of expression of the same genes in a normal or healthy subject.

## EXAMPLE 2

*Validation of gene expression profiling results using tissue microarrays*

Each of the transcripts identified as being differentially-expressed specifically in ovarian cancer was then further analysed using *in situ* hybridization or immunohistochemical staining of tissue microarrays constructed from a large cohort of primary ovarian tumor tissue. Such analysis confirms upregulation, down-regulation or total loss of expression of the transcripts identified in the microarray analysis of tumor samples.

Furthermore, as each of the samples in the tissue microarray have been clinicopathologically characterized (for example to identify cancer grade and/or disease stage) and the subjects from whom the tumors were isolated continuously monitored (to detect for example, death or relapse of cancer), changes with gene expression were also analysed for correlation with such parameters in order to determine predictive changes in gene expression.

The relative intensity and percentage of cells staining was determined and evaluated for associations with clinical stage and grade of disease and disease relapse using the Kaplan Meier method and log-rank test, and by univariate and bivariate analyses in a Cox proportional hazards model for gene expression and other clinical and pathologic predictors of outcome to determine the potential independent prognostic value of the markers being assessed.

Immunohistochemical analysis has been performed on several genes identified in gene profiling analysis of ovarian cancer samples. For example, SOX17, Ep-CAM and claudin 3 were shown by gene profiling analysis to be specifically up-regulated in ovarian cancer compared to normal ovaries (Figure 1 and Figure 2). Using immunohistochemical analysis, it was determined that SOX17, Ep-CAM and claudin 3 are upregulated in serous cancer, mucinous cancer, endometrioid cancer and clear cell ovarian cancer.

Furthermore, immunohistochemical analysis has been used to analyse the expression of several other genes that are specifically upregulated in mucinous ovarian cancer. In particular the expression of LI-cadherin (cadherin 17), meprin alpha and Galectin 4 as detected using immunohistochemistry is shown in Figure 3. There was a significant increase in protein detected in the mucinous ovarian cancer samples compared to the normal ovary sample and serous ovarian cancer sample.

Immunohistochemical analysis was also performed to analyse the expression of three genes that are known to be upregulated in ovarian cancer (CA125, MUC-1 and E-cadherin) (Figures 1 and 2).

5

### EXAMPLE 3

#### *Identification of prognostic markers of ovarian cancer*

Using a classical survival analysis to mine expression profiling data several genes that are associated with poor patient outcome (ie death or cancer relapse) have been identified (Tables 2 and 3). Such genes have clinical utility as prognostic indicators of disease.

Using detailed clinicopathological and postoperative data on all of the 51 patients included in our transcriptional profiling studies, including details of biochemical (eg. rising serum CA-125) and/or clinical recurrence of disease and overall survival, expression profiles were correlates with clinical parameters.

A preliminary survival analysis was performed on the 33 serous cancers within this cohort. The median follow-up time for these patients was 25.5 months from the date of primary laparotomy to the date of last follow-up or the date of death, and 21 of these patients (66%) were deceased from causes related to their malignancy.

Preliminary analysis of the expression profiles of these tumors identified several potential gene clusters that were associated with an increased risk of biochemical and clinical recurrence and overall survival, including the *EDD* gene (SEQ ID NO: 63). Exemplary prognostic markers for detecting ovarian cancer are shown in Tables 1 and 3. Preferred markers are indicated in Table 3.

Using immunohistochemical analysis two genes have been confirmed to be upregulated in serous ovarian cancer. In particular, sFRP4, a negative signalling protein of the Wnt pathway, and SOCS3, a negative signaller of IL-6 induced signalling are specifically upregulated in serous ovarian cancer when compared to normal ovarian tissue (Figure 4A).

Furthermore, using clinical patient data and correlating this information with gene expression levels using a Cox proportional hazards model, it has been shown that high

expression of sFRP4 correlates with a poor outcome in patients (n=127) with serous ovarian cancer (p=0.0056) (Figure 4B).

#### EXAMPLE 4

##### 5        *Validation of gene expression profiling results using quantitative RT-PCR*

Candidate diagnostic genes are screened by quantitative RT-PCR against ovarian cancer cell lines to both validate the transcript profiling data (ie check their up- or down-regulation). Candidate diagnostic genes are screened using mRNA isolated from a panel of 9 ovarian tumour cell lines, (A2780, SKOV3, OVCAR-3, IGROV-1, CAOV3, OV-90, 10 SW626, TOV-21G and TOV-112D), in addition to several other tumour cell lines including lines derived from breast, prostate and colorectal tumours, and immortalised (non-transformed) human ovarian surface epithelial cells and a primary normal breast epithelial cell line (184).

15        Total RNA is isolated from the normal and tumour cell lines, reverse transcribed into cDNA and used as template in a quantitative PCR using a LightCycler system (Roche Diagnostics). The relative amount of each gene product is determined by comparison to a standard housekeeping gene (GAPDH).

#### 20        EXAMPLE 5

##### *Identification of Novel Genes for Diagnosis of Ovarian Cancer*

We identified candidate genes with diagnostic potential from our list of aberrantly regulated genes by applying the following selection procedure: genes with a good transcript profile and low p-value (ie highly significantly up- or down-regulated in ovarian 25 cancer, as determined in Example 1); and mapping to areas of the genome that have been shown to be amplified or lost in ovarian cancer. Accordingly, it is likely that these genes are involved in the development and progression of ovarian cancer (ie putative oncogenes and tumour suppressor genes). Additional parameters for analysis included known or putative function in oncogenesis (eg signal transduction, regulation of cellular proliferation, apoptosis etc); and association with other forms of other tumours. Genes 30 identified in this analysis are shown in Table 3.

One method for the diagnosis of cancer comprises detecting modified DNA shed by the developing tumour into the blood stream. This can include the detection of mutations in 35 both oncogenes and tumour suppressor genes involved in the development and progression of ovarian cancer. Furthermore, it has been recently shown that aberrant

methylation of tumour suppressor genes, specifically hypermethylation of their gene promoters, frequently accompanies gene silencing in cancers, and indeed in some cases appears to be the predominant mechanism of gene silencing.

- 5 Combined with the knowledge of tumour nucleic acids circulating in the blood that reflect the biological characteristics of a tumour, the detection of methylation-specific tumour suppressor gene signatures for any given tumour type has promise as a specific and sensitive molecular test for detecting and monitoring cancer. Aberrant methylation is a frequent epigenetic event in epithelial ovarian cancer and many candidate tumour
- 10 suppressor genes of epithelial ovarian cancer have been shown to be hypermethylated in epithelial ovarian cancer, such as, for example BRCA1.

- In particular, expression of the candidate tumor suppressor gene MCC, has been shown to be down-regulated in epithelial ovarian cancer compared to normal ovarian tissue.
- 15 MCC appears to be involved in critical cell growth regulatory processes and maps to a chromosomal region hypothesised as containing a tumor suppressor gene in ovarian cancer. Furthermore, we have identified a CpG island within the predicted promoter sequence of the MCC gene, a critical feature of genes that are subject to gene silencing by hypermethylation and a known characteristic of tumor suppressor genes. Taken
- 20 together these data strongly implicate MCC as a candidate tumor suppressor gene involved in epithelial ovarian cancer.

Table 1  
Genes having modified expression in subjects suffering from ovarian cancer

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
a. upregulated genes				
NM_002354	Hs.892:235	Ep-CAM; TACSTD1, tumor-associated calcium signal transducer 1; epithelial glycoprotein	Lymphocyte antigen, plasma membrane, tumor antigen. Member of the GA733 family. C-terminus-associated antigen expressed on most normal epithelial cells and gastrointestinal carcinomas and functions as a homotypic calcium-independent cell adhesion molecule. The antigen is being used as a target for immunotherapy treatment of human carcinomas.	0
BC006428	Hs.15093:210; Hs.290304:1	HSPC195, hypothetical protein HSPC195	Homo sapiens cDNA FLJ10920 fis, clone OVARC1000384-resourcer.	0
NM_017697	Hs.24743:94	FLJ20171, hypothetical protein FLJ20171	contains 3 RNA recognition motifs	0
AW419186	Hs.257924:13	FLJ13782, Hypothetical protein FLJ13782	weakly similar to a drosophila transcription factor	0
AW830088	Hs.76550:164	MAL2	Mal2 T-cell differentiation protein; found thru interaction with TPD52 which is overexpressed in breast cancer. 4 TM are involved in vesicle transport	0
NM_004360	Hs.194657:233	CDH1, cadherin 1, type 1, E-cadherin (epithelial)	Tumor suppressor. Ca <sup>2+</sup> -dependent glycoprotein, mediates cell-cell interactions in epithelial cells. Mutations correlated with gastric, breast, colorectal, thyroid and ovarian cancer. Loss of function thought to contribute to progression in cancer by increasing proliferation, invasion, and/or metastasis. The ectodomain of this protein mediates bacterial adhesion to mammalian cells and the cytoplasmic domain is required for internalization.	0
NM_003761	Hs.172684:89	VAMP8, vesicle-associated membrane protein 8 (endobrevin)	Early endosome, membrane fraction, non-selective vesicle docking, non-selective vesicle transport, protein complex assembly, synaptic vesicle. Member of a family involved in docking or fusion of synaptic vesicles. Associated with the perinuclear vesicular structures of the early endocytic compartment.	0
NM_004415	Hs.349499	DSP, desmoplakin (DPI, DPII)	Cell shape and cell size control, cell-cell adherens junction, epidermal differentiation, intermediate filament, structural constituent of cytoskeleton. Acts as a site of attachment for intermediate filaments in desmosomes (intercellular junction in vertebrate epithelial cells). Compound heterozygosity for non-sense and missense mutations underlies skin fragility/woolly hair syndrome.	0
NM_013230	Hs.286124:357; Hs.375108	CD24: CD24 antigen (small cell lung carcinoma cluster 4 antigen)	Plasma membrane, humoral defense mechanism. Cell surface antigen; glycosyl phosphatidylinositol (GPI)-linked glycoprotein that differentiates and activates granulocytes and B lymphocytes.	0

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
NM_003710	Hs.233950:84;Hs.182265:2;Hs.7771:1	SPINT1, serine protease inhibitor, Kunitz type 1. Hepatocyte growth factor activator inhibitor.	Extracellular, membrane fraction, serine protease inhibitor. Member of the Kunitz family of serine protease inhibitors. Hepatocyte growth factor activator inhibitor is a potent inhibitor specific for HGF activator and is thought to be involved in regulation of proteolytic activation of HGF in injured tissues. Function unknown	0
NM_153345	Hs.17568:16	FLJ95566, hypothetical protein	Function unknown	0.0001
NM_015238	Hs.21543:36	KIAA0869, KIAA0869 protein; KIBRA	Function unknown	0.0002
A1282759	Hs.242463:1	KRT8, keratin 8	Cell structure, Cytoskeletal. May form intermediate filaments; type II keratin, member of a family of structural proteins. Disruption of mechanisms that normally regulate keratin expression in vivo could be related to inflammatory and neoplastic pancreatic disorders (Casanova 1999).	0.0002
A1393742	Hs.195087:46	ERBB3, v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)	Transmembrane receptor protein tyrosine kinase, epidermal growth factor receptor, integral plasma membrane protein, protein amino acid phosphorylation. Member of the ERBB gene family of receptor tyrosine kinases, elevated levels in certain human mammary tumor cell lines. A receptor for heregulin, capable of mediating HGL-stimulated tyrosine phosphorylation of itself. Epidermal growth factor contains both positive and negative determinants for interaction with ErbB-2/ErbB-3 heterodimers (Stortefers 2002)	0.0002
AW957300	Hs.294142:167	ESTs, Weakly similar to CYL1_HUMAN CYLICIN 1 [H.sepiens]	Function unknown	0.0002
NM_012474; W70171	Hs.75939:33;Hs.170864:1	UMPK, uridine monophosphate kinase	Catalyzes the phosphorylation of uridine monophosphate to uridine diphosphate. First step in production of pyrimidine nucleoside triphosphates required for RNA and DNA synthesis. An allele of this gene may play a role in mediating nonhumoral immunity to Hemophilus influenzae type B.	0.0003
AA165082	Hs.146386:47;Hs.113919:3	MAP7, microtubule-associated protein 7	Establishment and/or maintenance of cell polarity, microtubule associated protein, microtubule cytoskeleton organization and biogenesis, structural molecule. Predominantly expressed in cells of epithelial origin. Involved in microtubule dynamics and cell polarization and differentiation. Stabilizes microtubules, and may modulate microtubule functions. Studies of the related mouse protein suggest an essential role in microtubule function required for spermatogenesis.	0.0004
AA284679	Hs.25640:264;Hs.5372:2	CLDN3, claudin 3	Integral plasma membrane protein, pathogenesis, tight junction, transmembrane receptor. Member of the claudin family of integral membrane proteins; receptor for Clostridium perfringens enterotoxin;	0.0004



Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
NM_004433	Hs.168098:170	ELF3, E74-like factor 3 (eis domain transcription factor, epithelial-specific)	Embryogenesis and morphogenesis, transcription co-activator, transcription factor, transcription from POU II promoter, ETS domain transcriptional activator; activates expression of epithelial cell specific genes.	0.0004
AW247252	Hs.75514:181	NP, nucleoside phosphorylase	DNA modification, nucleoside nucleoside nucleotide and nucleic acid metabolism, purine-nucleoside phosphorylase. Enzyme purine nucleoside phosphorylase together with adenosine deaminase (ADA) serves a key role in purine catabolism, referred to as the salvage pathway. Mutations in either enzyme result in a severe combined immunodeficiency (SCID).	0.0004
NM_015925	Hs.361379, Hs.95697:59,Hs.9364 8:1	LISCH7, Liver-specific bHLH-Zip transcription factor	LISCH protein	0.0004
NM_022454	Hs.97984:22	SOX17, SRY (sex determining region Y)-box 17	Likely ortholog of mouse SRY-box containing gene 17; alias SOX17	0.0005
A1124756	Hs.5337:191	IDH2, Isocitrate dehydrogenase 2 (NADP+), mitochondrial	Carbohydrate metabolism, mitochondrion	0.0006
NM_003064	Hs.313:273,Hs.29789 5:1	SPP1, secreted phosphoprotein 1 (osteopontin, bone sialoprotein 1, early T-lymphocyte activation 1)	Osteopontin (bone sialoprotein); bone and blood vessel extracellular matrix protein involved in calcification and atherosclerosis. Increased expression is associated with breast tumor metastasis (Urquidí 2002). Role in HCC, especially in cancer-stromal interactions (Gotoh 2002). Association between levels of a biomarker, osteopontin, and ovarian cancer suggest its clinical usefulness (Kim 2002).	0.0006
BE382756	Hs.169902:319,Hs.27 5408:1	SLC2A1, Solute carrier family 2 (facilitated glucose transporter), member 1	Glucose transporter, membrane fraction, SLC2A1/GLUT1 - facilitated glucose transporter. Glucose transporter is an integral membrane glycoprotein that is involved in transporting glucose into most cells. 12 TMs. Role in transport of glucose across the blood-brain barrier. Consistent marker of ovarian epithelial malignancy (Kallir 2002). Marker for discriminating hepatocellular carcinoma from other carcinomas (Zimmernan 2002).	0.0006
BE512730	Hs.65114:718,Hs.279 437:1	KRT18, keratin 18	Cell shape and cell size control, embryogenesis and morphogenesis, intermediate filament, structural constituent of cytoskeleton. Component of intermediate filaments; type I epidermal keratin, strongly similar to murine Endo B. Expressed in single layer epithelial tissues of the body. Mutations linked to cryptogenic cirrhosis.	0.0006

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
NM_001769	Hs.1244:227,Hs.2305 59:1,Hs.242020:1	CD9: CD9 antigen (p24)	Plasma membrane, integral plasma membrane protein. Member of the transmembrane 4 superfamily (TM4SF); may mediate platelet activation and aggregation. Cell surface glycoprotein that is known to complex with Integrins and other transmembrane 4 superfamily proteins.	0.0006
A1701805; NM_019027	Hs.95549:147,Hs.229 59:1	FLJ20273, RNA-binding protein	Contains four RNA recognition motifs (RRM, RBD, or RNP)	0.0007
NM_008103	Hs.2719:108,Hs.5445 1:1	WFDC2, WAP four-disulfide core domain 2	Endopeptidase inhibitor, extracellular space, proteolysis and peptidolysis, spermatogenesis. Epididymis-specific secreted protein; may have a role in sperm maturation; arelong to a family of extracellular proteinase inhibitors. Expressed in pulmonary epithelial cells, and also expressed in some ovarian cancers.	0.0009
U81961	Hs.438580	SCNN1A, sodium channel, nonvoltage-gated 1 alpha	Amiloride-sensitive sodium channel, excretion, integral plasma membrane protein, membrane fraction, sodium transport. Alpha subunit of the amiloride-sensitive epithelial sodium channel; functions in nonvoltage-gated channel	0.0009
X69699; NM_013952	Hs.73149:72,Hs.2130 08:1	PAX8, paired box gene 8	Histogenesis and organogenesis, embryogenesis and morphogenesis, thyroid-stimulating hormone receptor, transcription factor. Member of the paired domain family of nuclear transcription factors; are involved in the ribosome assembly, required for normal thyroid development. PAX genes play critical roles during fetal development and cancer growth.	0.0009
A1027643	Hs.120912:12	ESTs	Function unknown	0.001
AA173992	Hs.7958:28	ESTs	Function unknown	0.0011
AB018249	Hs.10458:10	SCYA16, small inducible cytokine subfamily A (Cys-Cys), member 16.	Antimicrobial humoral response (sensu Invertebrata), cell-cell signaling, chemokine chemotaxis. Cytokine A16; lymphocyte and monocyte chemoattractant.	0.0011
NM_014791	Hs.184339:27	MELK, likely ortholog of maternal embryonic leucine zipper kinase.	KIAA0175 gene product; serine/threonine protein kinase domain	0.0011
NM_030874	Hs.18272:81	SLC38A1, solute carrier family 38, member 1	amino acid transporter A1 (ATA1), likely ortholog of mouse N-system amino acid transporter protein NAT2.	0.0012
NM_005882	Hs.6527:201	GPR68, G protein-coupled receptor 58	cell adhesion, cell-cell signalling, G-protein linked receptor, integral plasma membrane protein, G-protein linked receptor protein signalling pathway. Member of the G protein-coupled receptor family; similar to secretin and calcitonin receptors. 7 transmembrane domains, a much-like domain and cysteine box in the N-terminal region. Expressed in range of tissues, highest levels in thyroid, selectively within the monolayer of cuboidal epithelial cells of the smaller, more actively secreting follicles of human thyroid. Differentially expressed in melanoma cell lines with different	0.0012

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
AI669760	Hs.188881:6.Hs.189354:1	ESTs	metastatic potential (Zendman et al' 1999).	0.0013
NM_001730	Hs.84728:127	KLF5, Kruppel-like factor 5 (intestinal)	dbEST Library Tissue Type restricted to prostate	0.0014
AI355761	Hs.242463:2	KRT8, keratin 8	RNA polymerase II transcription factor, transcription from Pol II promoter. Zinc finger transcriptional activator; localizes to the nucleus and binds the epidermal growth factor response element, binds GC boxes.	0.0014
BE019020	Hs.85838:171	SLC18A3, solute carrier family 18 (monocarboxylic acid transporters), member 3 (MCT3)	Cell structure, Cytoskeletal. May form intermediate filaments; type II keratin, member of a family of structural proteins. Disruption of mechanisms that normally regulate keratin expression in vivo could be related to inflammatory and neoplastic pancreatic disorders (Casanova 1999).	0.0015
NM_001307 NM_002266	Hs.278562:101 Hs.159557:394	CLDN7, claudin 7 KPNA2, karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	Integral plasma membrane protein, membrane fraction, monocarboxylic acid transport, monocarboxylic acid transporter, Member of monocarboxylate transporter family; may function as a transporter (MCT3).	0.0016 0.0016
AW176120	Hs.9061:77	MGC2477, hypothetical protein MGC2477	Integral membrane protein, tight junction. Similar to murine Cldn7; DNA metabolism, G2 phase of mitotic cell cycle, NLS-bearing substrate-nucleus import, cytoplasm, importin alpha-subunit, nuclear localization sequence binding, nucleoplasm, regulation of DNA recombination, spindle pole body and microtubule cycle (sensu Saccharomyces). Karyopherin alpha 2 (importin alpha 1); subunit of the NLS (nuclear localization signal) receptor. KPNA2 protein interacts with the NLSs of DNA helicase Q1 and SV40 T antigen and are involved in the nuclear transport of proteins. KPNA2 also may play a role in V(D)J recombination.	0.0016
BE265489	Hs.3123:49	LLGL2, lethal giant larvae (Drosophila) homolog 2	Cytoskeleton, structural molecule. May associate with nonmuscle myosin II heavy chain. cDNA source cancer cell lines. 57% ID to m.musculus 1920362A tumor suppressor gene mgi1	0.0016
BE279383	Hs.26557:77	PKP3, plakophilin 3	Cell adhesion, intercellular junction. Desmosomal plaque proteins are members of the 'armadillo-repeat' multigene family and have important functions in cytoskeleton/cell membrane interactions.	0.0016

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
J05581; NM_002458	Hs.89803:128.Hs.296789:1	MUC1, mucin 1, transmembrane	Integral plasma membrane protein. Cell surface mucin glycoprotein expressed by most glandular and ductal epithelial cells and some hematopoietic cell lineages. Alterations in glycosylation in epithelial cancer cells. Marker for hepatocellular carcinoma. MUC1 metabolic complex conserved in tumor-derived and normal epithelial cells. Expression predictor of surgical outcome in mass-forming intrahepatic cholangiocarcinoma. Tyrosine kinase c-Src constitutes a bridge between cystic fibrosis transmembrane regulator channel failure and MUC1 overexpression in cystic fibrosis.	0.0016
AA531276	Hs.59509:9	ESTs (unnamed protein product)	Function unknown	0.0017
AW187128	Hs.231934:3	ESTs; weakly similar to A57717 transcription factor EC2	Function unknown	0.0018
AW368226	Hs.87928:25.Hs.229840:1	Ets-related transcription factor, ESX, epithelium-restricted Ets protein ESX-not in Unigene, but found using resourcerer.	Embryogenesis and morphogenesis, transcription co-activator, transcription factor, transcription from Pol II promoter.	0.0021
AK000733	Hs.23900:82	RACGAP1, Rac GTPase activating protein 1	Strongly similar to murine Racgap1 GTPase-activating protein for rac. The plexin-B1/Rac interaction inhibits PAK activation and enhances Semaphorin 4D ligand binding	0.0024
NM_014738	Hs.81892:95	KIAA0101 gene product	Function unknown; no significant hits with Superfamily	0.0025
NM_014586	Hs.109437:17	HUNK, hormonally upregulated neu tumor-associated kinase	Developmental processes, protein serine/threonine kinase, signal transduction, protein kinase containing SNF-1 (fam of serine/threonine kinases) domain; progesterone and estradiol regulated. Similar to murine Hunk.	0.0025
A1885516	Hs.95612:31.Hs.251688:1	desmocollin type 2a, desmocollin 2, isoform Desc2b preproprotein; desmosomal glycoprotein II/III; desmocollin-3-not in Unigene, but found using resourcerer.	Cell adhesion, intercellular junction	0.0027
AW194426	Hs.20726:17	ESTs	Function unknown	0.0027
NM_001982	Hs.199067:83.Hs.167388:1	ERBB3, HER3 (c-erb-B3), v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)	Epidermal growth factor receptor, integral plasma membrane protein, protein amino acid phosphorylation. Member of the ERBB gene family of receptor tyrosine kinases, elevated levels in certain human mammary tumor cell lines. A receptor for heregulin, capable of mediating HGL-stimulated tyrosine phosphorylation of itself.	0.0028
NM_007019	Hs.893002:85	UBE2C, ubiquitin carrier protein E2-C	Ubiquitin-dependent protein degradation, degradation of cyclin, protein modification, positive control of cell proliferation. Subunit of a complex with ubiquitin ligase activity; complex that exhibits cyclin-selective ubiquitin ligase activity.	0.0031

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
BE184455	Hs.251754:128,Hs.245742:1	SLPI, secretory leukocyte protease inhibitor (antileukoprotease)	Plasma protein, proteinase inhibitor. Secreted inhibitor which protects epithelial tissues from serine proteases. Found in various secretions including seminal plasma, cervical mucus, and bronchial secretions, has affinity for trypsin, leukocyte elastase, and cathepsin G. Its inhibitory effect contributes to the immune response by protecting epithelial surfaces from attack by endogenous proteolytic enzymes; the protein is also thought to have broad-spectrum anti-biotic activity.	0.0034
Y00815; NM_002840	Hs.75216:282,Hs.228792:1,Hs.245083:1	PTPRF, protein tyrosine phosphatase, receptor type, F	Cell adhesion, integral plasma membrane protein, transmembrane receptor protein, tyrosine phosphatase signaling pathway. Receptor-type protein tyrosine phosphatase F; interacts with the insulin receptor; has Ig-like and FN-III repeats in the extracellular domain	0.0035
AA708017	Hs.119944:14	ESTs	Function unknown	0.0038
AA256841	Hs.236884:24	ESTs, Highly similar to S02392 alpha-2-macroglobulin receptor precursor	Function unknown	0.0041
AW055308	Hs.31803:15	ESTs, Weakly similar to TRHY_HUMAN TRICHOHYALIN [H.sapiens]	Function unknown	0.0043
AI301658	Hs.290801:35, Hs.336228	EST	Function unknown	0.0044
T18997	Hs.180372:119; Hs.394609	BC12-like 1, Homo sapiens cDNA FLJ20750 fis, clone HEP05174 (hypothetical protein)	Function unknown	0.0044
AI798863	Hs.87191:8	ESTs	Function unknown	0.0049
JO3258	Hs.2062:146	VDR, vitamin D (1,25-dihydroxyvitamin D3) receptor	DNA binding, signal transduction, vitamin D3 receptor. Zinc-finger DNA-binding transcription factor. Genetic polymorphism determines bone mineral density. Stat1-vitamin D receptor interactions antagonize 1,25-dihydroxyvitamin D transcriptional activity and enhance stat1-mediated transcription.	0.0049
AA151647	Hs.68877:141,Hs.228686:1	CYBA, cytochrome b-245, alpha polypeptide	Cytochrome b, membrane, mitochondrion, superoxide metabolism. Alpha-subunit of cytochrome b245, primary component of the mitochondrial oxidase system of phagocytes. CYBA deficiency is associated with chronic granulomatous disease (CGD).	0.005

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
A1538613	Hs.135657:8	TMPPRS3 Transmembrane protease, serine 3	Integral membrane protein, proteolysis and peptidolysis. Contains a serine protease domain, a transmembrane domain, a LDL receptor-like domain, and a scavenger receptor cysteine-rich domain. Serine proteases are known to be involved in a variety of biological processes, whose malfunction often leads to human diseases and disorders. Expressed in fetal cochlea and many other tissues, and is thought to be involved in the development and maintenance of the inner ear or the contents of the perilymph and endolymph. Missense mutations in autosomal recessive sensorineural deafness. Identified as a tumor associated gene that is overexpressed in ovarian tumors.	0.0051
NM_018000	Hs.79741:18	FLJ10116, hypothetical protein FLJ10116	Function unknown	0.0051
NM_144724 AJ278016	Hs.124740:18 Hs.55566:35	hypothetical protein FLJ30532 ANKRD3, ankyrin repeat domain 3	59% identity to human Zinc finger protein 91	0.0051
NM_013994	Hs.76562:147	DDR1, discoidin domain receptor family, member 1	ATP binding, protein amino acid phosphorylation, protein binding, protein serine/threonine kinase.	0.0055
T09997: NM_001312	Hs.70327:186, Hs.211 478:1	CRIP-2, cysteine-rich protein 2	Cell adhesion, Integral plasma membrane protein, transmembrane receptor, protein tyrosine kinase. Epithelial-specific receptor protein tyrosine kinase; are involved in cell adhesion; has putative discoidin motifs in extracellular domain. DDR1 (CD167a) is a RTK that is widely expressed in normal and transformed epithelial cells and is activated by various types of collagen.	0.0055
BE302796	Hs.105097:115	TK1, thymidine kinase 1, soluble	Zn-finger LIM domain protein;208-amino acid protein containing 2 LIM domains	0.0055
NM_001067	Hs.158348:184, Hs.27 0810:2	TOP2A, topoisomerase (DNA) II alpha (170kD)	Cytoplasm, thymidine kinase. Generates thymidylate for DNA synthesis. TK1 gene expression together with TS, TP and DPD gene expression may play important roles in influencing the malignant behavior of epithelial ovarian cancer (Fujiwaki R 2002).	0.006
			DNA binding, DNA topoisomerase (ATP-hydrolyzing), nucleus. DNA topoisomerase II alpha; may relax DNA torsion upon replication or transcription. Involved in processes such as chromosome condensation, chromatid separation, and the relief of torsional stress that occurs during DNA transcription and replication. Catalyzes the transient breaking and rejoining of two strands of duplex DNA. The gene encoding this enzyme functions as the target for several anticancer agents and a variety of mutations in this gene have been associated with the development of drug resistance. Reduced activity of this enzyme may also play a role in ataxia-telangiectasia.	0.006

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
U48455	Hs.252189:148,Hs.248217:1	SDC4, syndecan 4 (amphiglycan, ryudocan)	Integral plasma membrane, proteoglycan syndecan. Syndecans are transmembrane heparan sulfate proteoglycans that appear to act as receptors or coreceptors involved in intracellular communication. Members of the MYC gene family and 4 members of the syndecan gene family are closely situated on 4 different chromosomes. Function unknown	0.0061
M79141	Hs.13234:39	ESTs	Function unknown	0.0062
A1955040	Hs.301684:5,Hs.265398:3	ESTs, Moderately similar to hypothetical protein FLJ20378 [Homo sapiens] [H.sapiens]	Function unknown	0.0065
NM_005560	Hs.11869:81,Hs.231010:1	LAMA5, laminin, alpha 5	Basement lamina, structural molecule. Widely expressed in adult tissues, with highest levels in lung, heart, and kidney. Fifth member of the alpha subfamily of vertebrate laminin chains. Possible basement membrane protein; contains laminin EGF-like domain, two extracellular laminin G domains. Cell-cell signaling, cytoplasm, extracellular space, protein binding. Protein that is induced by interferon.	0.0068
BE563085	Hs.833:97	ISG15, Interferon-stimulated protein, 15 kDa	Blood group antigen, cell adhesion, integral plasma membrane protein, signal transduction, transmembrane receptor. Lutheran blood group glycoprotein; may play role in cell-cell, cell-matrix adhesion, signal transduction; member of the Ig superfamily, has integrin-binding motifs, SH3 domains. Amiloride-sensitive sodium channel (weakly similar to Mus musculus PDZ domain actin binding protein)	0.0069
BE278288	Hs.155048:119	LU, Lutheran blood group (Auburger b antigen included)	Endoplasmic reticulum lumen, protein secretion. Strongly similar to rat Rn.4070 (CABP2); may bind calcium.	0.0074
NM_020859	Hs.278628:52	ShrmL, Shroom-related protein (KIAA1481 protein)	Member 8 of the interferon regulatory factor transcription factor family; has low similarity to IRF4, which is a lymphocytic transcription factor that stimulates B cell proliferation. Expressed in activated T/LAK lymphocytes	0.008
A1262789	Hs.93659:52	ERP70, protein disulfide isomerase related protein (calcium-binding protein, intestinal-related)	Strongly similar to murine Hn1	0.0082
NM_006147	Hs.11801:77	IRF8, Interferon regulatory factor 6	B link shows some homology to KIAA1294 but no known function	0.0087
R61463	Hs.16165:50	LAK-4P, expressed in activated T/LAK lymphocytes		0.009
A1878857; NM_016185	Hs.109706:285	HN1, hematological and neurological expressed 1 protein		
AK001763	Hs.73239:37	FLJ10901, hypothetical protein FLJ10901		

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
AC004770	Hs.4756:99	FEN1, flap structure-specific endonuclease 1	DNA repair enzyme, DNA replication, UV protection, double-strand break repair, double-stranded DNA binding, double-stranded DNA specific exodeoxyribonuclease, endonuclease, fatty acid desaturation, membrane fraction. Removes 5' overhanging flaps in DNA repair and processes the 5' ends of Okazaki fragments in lagging strand DNA synthesis.	0.0093
AI567421	Hs.273330:137	AGRN: agrin	AgRn is a neuronal aggregating factor that induces the aggregation of acetylcholine receptors and other postsynaptic proteins on muscle fibers and is crucial for the formation of the neuromuscular junction. Acts at the nerve-muscle synapse in the glomerular basal membrane and on T-lymphocytes.	0.0093
AW161386	Hs.13561:49	MGC4692: hypothetical protein MGC4692	Function unknown	0.0103
M85430	Hs.155181:546	VIL2, villin 2 (ezrin)	Cytoskeletal anchoring, microvillus. Regulates cell adhesion and cortical morphogenesis. The cytoplasmic peripheral membrane protein encoded by this gene functions as a protein-tyrosine kinase substrate in microvilli. As a member of the ERM protein family, this protein serves as an intermediate between the plasma membrane and the actin cytoskeleton. It plays a key role in cell surface structure adhesion, migration, and organization.	0.0108
AW250380	Hs.109059:124, Hs.24756:11	MRPL12, mitochondrial ribosomal protein L12	Protein synthesis, General cellular role, Ribosomal subunit, Mitochondrial, RNA-binding protein, Ribosome-associated.	0.0114
AI733848; NM_021220	Hs.71935:13	ZNF339, zinc finger protein 339	Zinc finger protein	0.0115
AF111856; NM_008424	Hs.105039:48	SLC34A2, solute carrier family 34 (sodium phosphate), member 2	SLC34A2: solute carrier family 34 (sodium phosphate), member 2; contains 8 predicted TMs and a cysteine-rich N-terminal region. Type 2 sodium-dependent phosphate transporter. member of the renal type II co-transporter family.	0.0121
BE386983; NM_138410	Hs.343214	CKLFSF7, chemokine-like factor super family 7	chemokine-like factor gene superfamily; transmb 4 superfamily	0.0131
AA433988	Hs.98502:8	MUC16, mucin 16, CA125	Mucin 16. Alias CA125 ovarian cancer antigen	0.0137
AW248314	Hs.9622:83	MRPS18A, mitochondrial ribosomal protein S18A	Mitochondrial small ribosomal subunit, protein biosynthesis, structural constituent of ribosome/ribosomal mitochondrial protein S18A	0.0149



Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
AA454501	Hs.43666:65	PTP4A3, protein tyrosine phosphatase type IVA, member 3	Phosphatase. PTPs are cell signaling molecules that play regulatory roles in a variety of cellular processes. Strong similarity to murine Ptp4a3 (Mm.4124). Overexpression of this gene in mammalian cells was reported to inhibit angiotensin-II induced cell calcium mobilization and promote cell growth. PRL3 (PTP4A3) expressed at high levels cancer metastases (Saha et al. 2001). PRL3 gene is important for colorectal cancer metastasis.	0.016
U33446	Hs.75799:116	PRSS8, protease, serine, 8 (proctasin)	Extracellular space, plasma membrane, serine type peptidase. A trypsinogen, member of the trypsin family of serine proteases. Highly expressed in prostate epithelia, one of several proteolytic enzymes found in seminal fluid. Protease-mediated regulation of sodium absorption is a function of human airway epithelia, and proctasin is a likely candidate for this activity.	0.0166
X98654	Hs.93837:43	PITPNM, phosphatidylinositol transfer protein, membrane-associated	Brain development, lipid metabolism, membrane fraction, phosphatidylinositol transporter, phototransduction. Catalyzes the transfer of phosphatidylinositol between membranes; similar to Drosophila rdgB.	0.0167
AI680149	Hs.44865:39,Hs.300819:19,Hs.293904:14	LEF1, Lymphoid enhancer-binding factor-1	Very strongly similar to murine Left1; may act as a transcription factor. Expressed in pre-B and T cells. Binds to T-cell receptor-alpha enhancer and confers maximal enhancer activity. A target gene ectopically activated in colon cancer, from selective activation of a promoter for a full-length LEF1 isoform that binds beta-catenin (HOVANES 2001).	0.0172
AF098158; NM_012112	Hs.9328:152	C20orf1, chromosome 20 open reading frame 1	ATP binding, GTP binding, cell proliferation, mitosis, nucleus spindle. Proliferation-associated nuclear protein; associates with the spindle pole and mitotic spindle during mitosis	0.0183
AB014551	Hs.155120:101, Hs.337774	ARHGEF2, rho/rac guanine nucleotide exchange factor (GEF) 2	Cell shape and cell size control, cell surface receptor linked signal transduction, guanyl-nucleotide exchange factor, microtubule cytoskeleton. Rho GTPases play a fundamental role in numerous cellular processes that are initiated by extracellular stimuli that work through G protein coupled receptors. The encoded protein may form complex with G proteins and stimulate Rho-dependent signals. Rho/Rac guanine nucleotide exchange factor (GEF) 2; associates with microtubules, stimulates GTP binding on Rac and Rho	0.0206
AI278023	Hs.89986:24,Hs.290780:1	ESTs	Function unknown	0.0208

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
Z95152	Hs.178695:25,Hs.79107:1	MAPK13, mitogen-activated protein kinase 13	MAP kinase, antimicrobial humoral response (sensu Invertebrata), cell surface receptor, signal transduction, chemotaxis, stress response. MAP kinases act as an integration point for multiple biochemical signals, and are involved in a wide variety of cellular processes such as proliferation, differentiation, transcription regulation and development. Are activated by proinflammatory cytokines and cellular stress. Transcription factor ATF2, and microtubule dynamics regulator stathmin are substrates of this kinase.	0.0217
AW840171	Hs.265388:7	ESTs, Moderately similar to hypothetical protein FLJ20378 [Homo sapiens] [H-sapiens]	Function unknown	0.0222
D49441	Hs.155981:53	MSLN, mesothelin	Cell adhesion, cell surface antigen, membrane. <i>Pre-pro-megakaryocyte</i> potentiating factor. An antibody that reacts with ovarian cancers and mesotheliomas was used to isolate a cell surface antigen named mesothelin. Although the function of mesothelin is unknown, it may play a role in cellular adhesion and is present on mesothelium, mesotheliomas, and ovarian cancers.	0.0225
AW797437	Hs.69771:282,Hs.444:1,Hs.294163:1	EST, CM1-UM0039-030400-173-a09	Function unknown	0.0229
BE396290	Hs.5097:261	SYNGR2, synaptogyrin 2	Integral plasma membrane protein, member of a family of transmembrane synaptic vesicle proteins, specialized secretory organelles that store neurotransmitters in nerve terminals, and release them by fusing with the presynaptic plasma membrane during exocytosis.	0.0229
A1656166; NIM_025080	Hs.7331	ASRGL1; asparaginase like 1	glycoprotein catabolism	0.02
NIM_002145	Hs.2733:25	HOXB2, homeo box B2, Hox2H protein	Circulation, developmental processes, transcription factor. Member of homeodomain family of DNA binding proteins; may regulate gene expression, morphogenesis, and differentiation. Genes of the HOXB (or HOX2) complex are expressed specifically in erythromegakaryocytic cell lines, some are expressed only in hematopoietic progenitors.	0.024
AW959311	Hs.87019:8; Hs.172012	Hypothetical protein DKFZp434J037	probable serine/threonine protein kinase; KIAA0537	0.0251

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
NM_000269	Hs.118638:168,Hs.276104:1,Hs.276127:1,Hs.276248:1	NME1, non-metastatic cells 1, protein (NM23A)	Transcription factor and nucleoside diphosphate kinase; has a role in the transcriptional regulation of c-myc expression. Mutations in NME1 have been identified in aggressive neuroblastomas.	0.0257
AA379597	Hs.5189:87,Hs.277192:1	HSPC150, HSPC150 protein similar to ubiquitin-conjugating enzyme	Similar to ubiquitin conjugating enzyme	0.0259
BE148235	Hs.193083:100	Homo sapiens cDNA FLJ14201 fis, clone NT2RP3002955	high homology to ARP-3 actin-like protein	0.0259
AI683243; AI587638	Hs.97258	ESTs	Mod similarity to S29539 ribosomal protein L13a	0.03
AF111713	Hs.286218:64	JAM1, junctional adhesion molecule	Cell motility, inflammatory response, intercellular junction. Role in the regulation of tight junction assembly in epithelia. Ligand of JAM is required for reovirus-induced activation of NF-kappa-B and apoptosis. Role in lymphocyte homing.	0.0261
BE381635	Hs.75725:450,Hs.274751:1,Hs.277482:1,Hs.277488:1	TAGLN2, transgelin 2	Complex assembly protein. Homolog of the protein transgelin, which is one of the earliest markers of differentiated smooth muscle. Function not yet determined. Are an actin-binding protein.	0.0276
D14697	Hs.77393:201,Hs.247769:1	FOPS, farnesyl diphosphate synthase (farnesyl pyrophosphate synthetase, dimethylallyltransferase, geranyltransferase)	Farnesyl pyrophosphate synthetase (farnesyl diphosphate synthase); part of the cholesterol synthesis pathway.	0.0278
AW194384	Hs.94814	MGC2865, Hypothetical protein MGC2865	Function unknown.	0.0295
T47384	Hs.278613:145	IF127, interferon, alpha-inducible protein 27	Integral membrane protein. Isolated from estradiol-treated human breast carcinoma cells. Induced by interferon-alpha in human cell lines of different origin, expression is independent of the presence of estradiol receptor in the cells.	0.03
U17760	Hs.301103:71,Hs.76517:24,Hs.198068:1	LAMB3, Laminin, beta 3 (lncaln (125kD), kalinin (140kD), BM600 (125kD)) (Acn NM_000228)	Epidermal differentiation, laminin-5, structural molecule. Member of a family of basement membrane proteins. LAMB3 serves as the beta chain in laminin-5. Mutations in LAMB3 have been identified as the cause of various types of epidermolysis bullosa.	0.0304
AU078517	Hs.184276:142	SLC9A3R1, solute carrier family 9 (sodium/hydrogen exchanger), isoform 3 regulatory factor 1	Actin cytoskeleton, protein complex assembly. Regulatory cofactor of the NHE3 (SLC9A3) sodium/hydrogen antiporter. Interacts with merlin (NF2) and ERM family members; has two PDZ domains. Structural determinants in interaction of beta 2 adrenergic and platelet-derived growth factor receptors	0.0312

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
AW880841	Hs.96908 Hs.74427:112	PIG11, p53-induced protein	Negative control of cell proliferation, stress response. May generate or respond to oxidative stress, may have a role in p53-dependent apoptosis Polyak K, Xia Y, Zweier JL, Kinzler KW, Vogelstein B. A model for p53-induced apoptosis. Nature. 1997 Sep 18;389(6848):300-5.	0.0314
H24185 BE814410	Hs.92818:91 Hs.23044:51	BM-009, hypothetical protein BM-009 MGC18386, hypothetical protein, similar to RIKEN cDNA	Function unknown Function unknown.	0.0314 0.0326
H18423	Hs.82685:37	CD47: CD47 antigen (Rb-related antigen, Integrin-associated signal transducer)	Oncogenesis, plasma membrane, plasma glycoprotein, cell-cell matrix adhesion, Integral plasma membrane proteoglycan, Integrin receptor signal signalling pathway. Similar to Rb-antigen; may interact with Integrins and have a role in intracellular calcium increase during cell adhesion.	0.0336
AU076611; NM_006636	Hs.154672:123	MTHFD2, methylene tetrahydrofolate dehydrogenase (NADP+ dependent); methenyltetrahydrofolate cyclohydrolase	Electron transporter, methenyltetrahydrofolate cyclohydrolase, mitochondrial. encodes a nuclear-encoded mitochondrial bifunctional enzyme with methenyltetrahydrofolate dehydrogenase and methenyltetrahydrofolate cyclohydrolase activities. may provide formyltetrahydrofolate for formylmethionyl tRNA synthesis; involved in initiation of mitochondrial protein synthesis.	0.0342
A859390	Hs.288940:49	TMEM8, five-span transmembrane protein M83; type I transmembrane protein	Integral plasma membrane protein. Type I transmembrane protein; contains five membrane-spanning domains	0.0345
AA159216	Hs.55505:57	FLJ20442, hypothetical protein FLJ20442	Contains a dual specificity protein phosphatase catalytic domain; 34% similar to protein-tyrosine phosphatase	0.0354
AF119865; NM_021129	Hs.184011:156	PP, pyrophosphatase (inorganic)	Inorganic diphosphatase, phosphate metabolism. Catalyzes the hydrolysis of pyrophosphate to inorganic phosphate	0.0358
BE513613; NM_005720	Hs.11538:275	ARPC1B, actin related protein 2/3 complex, subunit 1A (41 kD)	Cell motility, structural constituent of cytoskeleton. Arp2/3 complex, subunit 1A; involved in assembly of the actin cytoskeleton, may have a role in protrusion of lamellipodia	0.0367
NM_012153	Hs.182339	EHF: ets homologous factor	DNA binding, tumor suppressor, cell proliferation, developmental processes, transcription activating factor. Member of the ESE subfamily of Ets transcription factors	0.0404
AW772298	Hs.21103:40;Hs.2687 84:2;Hs.102950:1	Homo sapiens mRNA; cDNA DKFZp564B076 (from clone DKFZp564B076)	Alias coat protein gamma-cop	0.0423
H16546	Hs.118668:66	PP591, hypothetical protein PP591	Hypothetical protein PP591 (Novel Human cDNA clones with function of inhibiting cancer cell growth; unpublished)	0.043

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
AA279661	Hs.83753:244,Hs.301236:3	SNRPB, small nuclear ribonucleoprotein polypeptides B and B1	Spliceosome, mRNA splicing, small nuclear ribonucleoprotein. U1 and U2 snRNP protein; component of snRNP complexes, required units of the spliceosome	0.0446
BE001596	Hs.85266:102	ITGB4, Integrin, beta 4	Cell adhesion receptor, Integrin, invasive growth, oncogenesis. Beta 4 subunit of Integrin; involved in cell-cell and cell-matrix interactions; member of a family of cell-surface proteins. Binding of beta 4 to plectin is essential for the proper formation and function of hemidesmosomes.	0.0453
BE246444	Hs.283685:148,Hs.232028:2	FLJ20396, hypothetical protein FLJ20396	100%/175aa unnamed protein g7020468	0.0453
X54942	Hs.83758:34	CKS2, CDC28 protein kinase 2	Cell proliferation, regulation of CDK activity. Similar to S. pombe p13suc1; binds and regulates CDK-cyclin complexes, expressed in different patterns through the cell cycle in HeLa cells, which reflects specialized role for the encoded protein.	0.0478
AA305599	Hs.238205:36	PRO2013, hypothetical protein PRO2013	Function unknown	0.0483
AF019228	Hs.8038:84	RAB3D, member RAS oncogene family	RAB small monomeric GTPase, hemocyte development. GTP-binding protein; are involved in vesicle transport; member of the RAB family of small GTPases. Alias GOV, that is overexpressed in glioblastoma multiforme tissue as compared to normal brain tissue. GOV is also highly expressed in recurrent glioma, colon tumor metastatic to brain, breast tumors, prostate tumors, and several tumor cell lines	0.0485
NM_001949	Hs.1189:65,Hs.286939:2	E2F3, E2F transcription factor 3	Protein binding, transcription factor, transcription initiation from Pol II promoter. Involved in cell cycle regulation, binds retinoblastoma protein (Rb). E2F family plays a crucial role in the control of cell cycle and action of tumor suppressor proteins and is also a target of the transforming proteins of small DNA tumor viruses.	0.049
AF217513	Hs.279905:73,Hs.283649:4	ANKT, nucleolar protein ANKT	clone HQ0310 PRO0310p1 nucleolar protein ANKT - no functional data	0.0504
AW513143	Hs.98367:8	ESTs	Expressed in uterus	0.0535
AJ245671	Hs.12644:73	EGFL6, EGF-like-domain; multiple 6	Cell cycle, oncogenesis, Integrin ligand, extracellular space. Member of the epidermal growth factor (EGF) repeat superfamily; contains an EGF-like-domain. Expressed early during development, and its expression has been detected in lung and meningioma tumors.	0.0568
AA084248	Hs.85339:64	GPR39, G protein-coupled receptor 39	G-protein linked receptor, G-protein coupled receptor protein signaling pathway, integral plasma membrane protein.	0.19

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
U62801	Hs.79361:85	KLK6, kallikrein 6 (neurosin, zyme)	Serine type peptidase, pathogenesis. Neurosin (protease M, zyme); a serine protease that cleaves amyloid precursor protein (APP). Growing evidence suggests that many kallikreins are implicated in carcinogenesis and some have potential as novel cancer and other disease biomarkers.	0.0159
D49441	Hs.155981:53	MSLN, mesothelin	Cell adhesion, cell surface antigen, membrane. Pre-pro-megakaryocyte potentiating factor. An antibody that reacts with ovarian cancers and mesotheliomas was used to isolate a cell surface antigen named mesothelin. Although the function of mesothelin is unknown, it may play a role in cellular adhesion and is present on mesothelium, mesotheliomas, and ovarian cancers.	0.147
X51630	Hs.1145:22.Hs.29685 1:1	WT1, Wilms tumor 1	Nucleus, transcription factor, transcription regulation. 4 Zn finger domains. Functions in kidney and gonad proliferation and differentiation. Mutations in this gene are associated with the development of Wilms tumors in the kidney or with abnormalities of the genitourinary tract.	0.2938
AB018305	Hs.5378:149	SPON1, spondin 1, (s-spondin) extracellular matrix protein	Extracellular matrix protein. Very strongly similar to rat F-spondin (Rn.7548); may have a role in the growth and guidance of axons.	0.3394
AA433988	Hs.98502:8	MUC16, mucin 16, CA125	Mucin 16. Alias CA125 ovarian cancer antigen	0.6568
NM_006149	Hs.5302:132	LGALS4, lectin, galactoside-binding, soluble, 4 (galectin 4)	Lectin, cytosol, cell adhesion, plasma membrane. Binds to beta galactoside. Involved in cell adhesion, cell growth regulation, inflammation, immunomodulation, apoptosis and metastasis; member of a family of lectins. LGALS4 is an S-type lectin that is strongly underexpressed in colorectal cancer.	0.0001
AA315933	Hs.120878:17	Homo sapiens, clone MGC:32871 IMAGE:4733535, mRNA, complete cds	Function unknown	0.0001
U47732	Hs.84072:110	TM4SF3, transmembrane 4 superfamily member 3	Integral plasma membrane protein, lysosome, pathogenesis, protein amino acid glycosylation, signal transducer, tumor antigen. Cell surface glycoprotein defined by the monoclonal antibody CO-029 is a 27- to 34-kD membrane protein expressed in gastric, colon, rectal, and pancreatic carcinomas but not in most normal tissues	0.0028

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
NM_005588	Hs.179704	MEP1A, meprin A alpha, PABA peptide hydrolase	metalloprotease located apically and secreted by epithelial cells in normal colon; degrades broad range of ECM components in vitro; proposed role in tumour progression by facilitating migration, intravasation and metastasis	0.01
AW503395	Hs.5541:12	ATP2A3, ATPase, Ca++ transporting, ubiquitous	Endoplasmic reticulum, adenosinetriphosphatase, small molecule transport, calcium-transporting ATPase, integral plasma membrane protein. Sarco/endoplasmic reticulum Ca2+-ATPase; pumps calcium.	0.0154
NM_004063	Hs.89436:50	CDH17, cadherin 17, LI cadherin (liver-intestine)	Cell adhesion, integral plasma membrane protein, membrane fraction, small molecule transport, transporter. Member of the cadherin family of calcium-dependent glycoproteins; facilitates uptake of peptide-based drugs, may mediate cell-cell interactions. Component of the gastrointestinal tract and pancreatic ducts. Intestinal proton-dependent peptide transporter in the first step in oral absorption of many medically important peptide-based drugs.	0.0172
A073913	Hs.100686:20	LOC155465, anterior gradient protein 3	Oncogenesis	0.0266
A928445	Hs.92254:80	SYTL2: synaptotagmin-like 2	Synaptotagmin-like protein of the C2 domain-containing family of proteins. Although the specific function of the synaptotagmin-like proteins is unknown, a role in regulation of synaptic vesicle trafficking via their C2 domains has been suggested. Region of weak similarity to murine Gph	0.08
W40460	Hs.144442:5	PLA2G10: phospholipase A2, group X	Extracellular, secreted phospholipase A2. Group X secretory phospholipase_a2; hydrolyzes the phospholipid sn-2 ester bond; member of the phospholipase family	0.1888
AA132981	Hs.212533:4	Homo sapiens cDNA: FLJ22572 fis, clone HSI02313	Function unknown	0.1985
AF111856	Hs.105039:48	SLC34A2, solute carrier family 34 (sodium phosphate), member 2	SLC34A2: solute carrier family 34 (sodium phosphate), member 2; contains 8 predicted TMs and a cysteine-rich N-terminal region. Type 2 sodium-dependent phosphate transporter. member of the renal type II co-transporter family.	0.5078
AA143854	z085a02.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone IMAGE:591722 5', mRNA sequence		Function unknown	0.036

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
<b>b. prognostic indicators</b>				
AA046217	Hs.105370:2	ESTs	Function unknown	0.00
NM_015902		EDD:Homo sapiens progesterin induced protein (DD5), mRNA, VERSION NM_020967.1 GI	Soluble fraction, cell proliferation, ubiquitin-protein ligase, ubiquitin conjugating enzyme, ubiquitin-dependent protein degradation. Member of the HECT family of proteins; may function as an E3 ubiquitin-protein ligase. This gene is localized to chromosome 8q22, a locus disrupted in a variety of cancers. This gene potentially has a role in regulation of cell proliferation or differentiation.	0.00
T83882	Hs.97927:20	ESTs	Function unknown	0.01
#(NOCAT)		NM_001615:Homo sapiens actin, gamma 2, smooth muscle, enteric (ACTG2), mRNA, variant 1, mRNA.	Structural protein of muscle. Gamma 2 actin; enteric-type, smooth muscle cell actin.	0.01
AB040888		Homo sapiens mRNA for KIAA1455 protein, partial cds	Function unknown	0.01
AA628980	Hs.192371:3	DSCR8	Function unknown	0.01
A623351	Hs.172148:51	down syndrome critical region protein DSCR8 ESTs	Function unknown	0.01
AW614420	Hs.204354:383	ARHB	Function unknown	0.01
		ras homolog gene family, member B	RHO small monomeric GTPase, RHO protein signal transduction, peripheral plasma membrane protein. Ras-related GTP binding protein of the rho subfamily, member B; may regulate assembly of actin stress fibers and focal adhesions; very strongly similar to murine Arhb.	0.01
AA243499	Hs.104800:23	hypothetical protein FLJ10134	Highly similar to murine p19.5; are a membrane protein	0.01
AF251237	Hs.112208:16	GAGED2 XAGE-1 protein	GAGE genes are expressed in a variety of tumors and in some fetal and reproductive tissues. This gene is strongly expressed in Ewing's sarcoma, alveolar rhabdomyosarcoma and normal testis. The protein encoded by this gene contains a nuclear localization signal and shares a sequence similarity with other GAGE/PAGE proteins. Because of the expression pattern and the sequence similarity, this protein also belongs to a family of CT (cancer-testis) antigens.	0.01
A1970797	Hs.64859:16	ESTs	Function unknown	0.01
AF145713	Hs.61490:51	SCHIP1 schwannomin-interacting protein 1	Cytoplasm. Associates with the neurofibromatosis type 2 protein schwannomin (NF2); contains a coiled-coil domain Proteome	0.01



Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
X78565	Hs.289114:173,Hs.74637:1	TNC hexabrachion (tenascin C, cytactin)	Cell adhesion, extracellular matrix, cell adhesion receptor, ligand binding or carrier. Hexabrachion (tenascin C), an extracellular matrix glycoprotein; has epidermal growth factor-like repeats	0.01
T97307		gb:ye53h05.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:121497 3', mRNA sequence.	Function unknown	0.01
BE243845	Hs.76511:418	CTGF connective tissue growth factor	Cell motility, plasma membrane, soluble fraction, response to wounding, extracellular matrix, extracellular space, epidermal differentiation, cell growth and maintenance, insulin-like growth factor binding, insulin-like growth factor receptor binding protein. Connective tissue growth factor; binds IGF, may have a role in regulating normal and neoplastic cell growth	0.01
AW088302	Hs.182183:214,Hs.325474:172,Hs.283080:7	CALD1 caldesmon 1	Cytoskeleton, actin binding, calmodulin binding, tropomyosin binding. Protein of unknown function. Actomyosin regulatory protein, non-muscle form	0.01
AL133561	Hs.241428:5	DKFZP434B081 protein	Function unknown	0.01
BE313555	Hs.7252:158	RAH17 retinoic acid induced 17	Function unknown	0.02
X07820	Hs.2258:1	MMP-10 matrix metalloproteinase 10 (MMP10; stromelysin 2)	Zinc binding, extracellular space, extracellular matrix, metalloendopeptidase, proteolysis and peptidolysis. Stromelysin 2; matrix metalloprotease that degrades connective tissue	0.02
AI973018	Hs.15725:77	IER5 immediate early response 5	Function unknown. A related mouse gene may play an important role in mediating the cellular response to mitogenic signals.	0.02
AF084545		Homo sapiens versican Vht Isoform, mRNA, partial cds	Function unknown	0.02
U41518	Hs.74602:148,Hs.767:1	AQP1 aquaporin 1 (channel-forming integral protein, 28kD)	Excretion, water transport, water transporter, integral plasma membrane protein. Aquaporin 1 (channel-forming integral protein); member of a family of water-transporters	0.02
Z11894		H. sapiens rearranged mRNA for immunoglobulin kappa chain (VNU)		0.02
AW138190	Hs.180248:8	ZNF124 zinc finger protein 124 (HZF-18)	DNA binding, C2H2 zinc-finger protein 124	0.02
BE088548	Hs.42346:83,Hs.6975:42	MYO22 myozenin 2	caldesmon-binding protein calsardin-1	0.02

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
W47186	Hs.166172:50	ARNT aryl hydrocarbon receptor nuclear translocator	Nucleus, transcription factor, transcription co-activator, transcription, DNA-dependent, protein-nucleus import, translocation, aryl hydrocarbon receptor nuclear translocator. Aryl hydrocarbon receptor nuclear translocator; used in receptor translocation from cytosol to nucleus	0.02
A1766870	Hs.54277:76	DXS9928E DNA segment on chromosome X (unique) 9928	Nucleus. Has many charged residues and a possible nuclear localization signal	0.02
X02761	Hs.287820:73;Hs.321592:1	FN1 fibronectin 1	Cell adhesion, cell motility, cell adhesion, soluble fraction, signal transduction, extracellular matrix, extracellular space. Fibronectin 1; member of family of proteins found in plasma and extracellular matrix	0.02
AW988613	Hs.79428:166	BNIP3 BCL2/adenovirus E1B 19kD-interacting protein 3	Anti-apoptosis, apoptosis inhibitor. Bcl2-related protein 3; binds antiapoptotic viral E1B 19 kDa protein and cellular Bcl2 protein	0.02
AW972565	Hs.32399:24	ESTs, Weakly similar to S51797 vasodilator-stimulated phosphoprotein [H.sapiens]	Function unknown	0.02
AF045229	Hs.82280:81	RGS10 regulator of G-protein signalling 10	Regulator of G protein signaling (RGS) family members are regulatory molecules that act as GTPase activating proteins (GAPs) for G alpha subunits of heterotrimeric G proteins. RGS proteins are able to deactivate G protein subunits of the G alpha, Gq alpha and Gq alpha subtypes. They drive G proteins into their inactive GDP-bound forms.	0.02
AW953853	Hs.292833:19	PAEP progesterone-associated endometrial protein (placental protein 14, pregnancy-associated endometrial alpha-2-globulin, alpha uterine protein)	Developmental processes. Placental protein 14 (Glycodelin); member of lipocalin superfamily, highly similar to beta-lactoglobulins	0.02
U52426	Hs.74597:75;Hs.157615:3	STIM1 stromal interaction molecule 1	Integral plasma membrane protein, positive control of cell proliferation. Very strongly similar to murine Stim1; are a transmembrane stromal cell protein	0.02
F06700	Hs.7878:115	IFRD1 interferon-related developmental regulator 1	Myoblast determination. Strongly similar to rat interferon-related developmental regulator 1; may play a role in muscle differentiation	0.02
A1788883	Hs.87181:3	ESTs C4001170;g1[8863178]gbjAAAF30402.1[AF109924.1 (AF109924) sulfatase 1 precursor [Helioma]	Function unknown	0.03
NA				0.03
H52761	Hs.141475:24	Homo sapiens cDNA clone IMAGE:178663	Function unknown	0.03

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
BE546947	Hs.44276:43	HOXC10 homeo box C10	Embryogenesis and morphogenesis, positive control of cell proliferation, RNA polymerase II transcription factor. Homeobox C10, member of the homeobox developmental regulator family; binds with HOXA13 and HOXC13 to the Lamin B2 origin; ortholog of Drosophila Abdominal-B	0.03
AU076643	Hs.313:257,Hs.32991 0:1	SPP1. secreted phosphoprotein 1 (osteopontin, bone sialoprotein I, early T-lymphocyte activation 1)	Ossification, extracellular matrix, skeletal development. Osteopontin (bone sialoprotein); bone and blood vessel extracellular matrix protein involved in calcification and atherosclerosis	0.03
#(NOCAT)		NIM_015902*:Homo sapiens progesterone induced protein (DD5), mRNA. VERSION NIM_020967.1 GI		0.03
U20536	Hs.3280:20	CASP8 caspase 8, apoptosis-related cysteine protease	Induction of apoptosis, cysteine-type peptidase, proteolysis and peptidolysis. Caspase 8; a cysteine (thiol) protease; related to the ICE-subfamily of caspases	0.03
AA581602		ESTs	Function unknown	0.03
AJ245210	Hs.41840:7	gb:Homo sapiens mRNA for immunoglobulin gamma heavy chain variable region, partial, clone 1A-4G21.	Function unknown	0.03
X65965		H.sapiens SOD-2 gene for manganese superoxide dismutase		0.03
AI808770	Hs.30258:9	ESTs	Function unknown	0.03
BE386490	Hs.279663:51	PIR P1rh	Nucleus, transcription co-factor, transcription from Pol II promoter. Putative cofactor of the NF/CTF1 transcriptional activator	0.03
AW581992	Hs.301434:104,Hs.32 9017:1	KIAA1387 KIAA1387 protein	Function unknown	0.03
U77534		Human clone 1A11 immunoglobulin variable region (VH5-D-JH4) gene, partial cds	Function unknown	0.03
AL034417	Hs.11169:194,Hs.109 58:1,Hs.74137:1	Gene 33/Mig-6	Function unknown	0.03
L10343	Hs.112341:96,Hs.196 8:1	Homo sapiens elafin precursor, gene, complete cds	Function unknown	0.03
AW518944	Hs.76325:80,Hs.2312 99:1	IGJ immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides	Linker protein for immunoglobulin alpha and mu polypeptides	0.03

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
W28729	Hs.236510:6	Human retina cDNA randomly primed sublibrary Homo sapiens cDNA, mRNA sequence	Function unknown	0.03
AI640160	Hs.74131:4	ARSE arylsulfatase E (chondrodysplasia punctata 1)	Arylsulfatase, skeletal development. Arylsulfatase E; likely involved in warfarin embryopathy.	0.03
U11862	Hs.75741:62	ABP1 amiloride binding protein 1 (amine oxidase (copper-containing))	Metabolism, peroxisome, amine oxidase, drug binding. Diamine oxidase (D-amine-acid oxidase histaminase, amiloride-binding protein); deaminates putrescine and histamine	0.03
AW295980	Hs.252741:3	ESTs	Function unknown	0.03
X59135	Hs.156110:4	H.sapiens mRNA for immunoglobulin 0-81VL		0.03
BE466173	Hs.378794	Homo sapiens mRNA; cDNA DKFZp686N0118 (from clone DKFZp686N0118)	Function unknown	0.03
#(NOCAT)		Target Exon		
AI354722	Hs.127216:24	hypothetical protein FLJ13465	Function unknown	0.03
M80484	Hs.169825:45,Hs.408 :1	Human collagen type IV alpha 5 chain (COL4A5) gene, 5' end	Function unknown	0.04
AA829286	Hs.332053:48,Hs.336 462:10	SAA1 serum amyloid A1	Inflammatory response, high-density lipoprotein. Member of the serum amyloid A protein family; member of high density apolipoproteins.	0.04
AI333771	Hs.82204:8,Hs.22836 3:1	ESTs	Function unknown	0.04
BE465867: NM_014992	Hs.197751:66	DAA1 dishevelled associated activator of morphogenesis 1	The protein encoded by this gene contains FH domains and belongs to a novel FH protein subfamily implicated in cell polarity, thought to function as a scaffolding protein.	0.04
BE516902	Hs.285313:145,Hs.40 55:43	COPEB core promoter element binding protein	A transcriptional activator, capable of activating transcription approximately 4-fold either on homologous or heterologous promoters. The DNA binding and transcriptional activity of this protein, in conjunction with its expression pattern, suggests that this protein may participate in the regulation and/or maintenance of the basal expression of pregnancy-specific glycoprotein gene and possibly other TATA box-less genes.	0.04

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
AA430373		gb:zv2011.51 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:789869 3' similar to gb:M63438 IG KAPPA CHAIN PRECURSOR V-III REGION (HUMAN);, mRNA sequence.	Function unknown	0.04
R27430	Hs.271565:3	ESTs	Function unknown	0.04
BE397335	Hs.283713:68	CTHRC1 collagen triple helix repeat containing 1 ESTs	Function unknown	0.04
AW284102	Hs.39168:16	Target Exon	Function unknown	0.04
NA		KIAA0591 protein	Function unknown	0.04
AW952323	Hs.129908:39	ESTs	Function unknown	0.04
AA088177	Hs.172870:13	ESTs	Function unknown	0.04
BE614567	Hs.19574:123	MGC5468 hypothetical protein MGC5468	Function unknown	0.04
AL079658	Hs.338207:139,Hs.146559:1	FRAP1 FK506 binding protein 12-rapamycin associated protein 1	DNA repair, DNA recombination, cell cycle control, 1'-phosphatidylinositol 3-kinase, inositol/phosphatidylinositol kinase, FKBP-rapamycin associated protein; phosphatidylinositol kinase that may mediate rapamycin inhibition of the cell cycle progression through G1	0.04
NM_002776	Hs.69423:46,Hs.275464:1	KLK10 kallikrein 10 (KLK10) (PRSSL1) (nes1)	Extracellular, serine-type peptidase. Putative serine protease	0.04
BE261944	Hs.118625:62	CYB561 cytochrome b-561	Energy pathways, secretory vesicle, cytochrome b5 reductase, secretory vesicle membrane, integral plasma membrane protein. Cytochrome b561; serves as a biological marker for adrenergic secretory vesicles	0.04
NM_006379	Hs.171921:50	SEMA3C sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3C	Drug resistance, immune response, cell growth and maintenance. Semaphorin E; member of a protein family involved in neuronal growth cone guidance	0.04
AI002238	Hs.11482:19	SFRS11 splicing factor, arginine/serine-rich 11	Nucleus, mRNA splicing, mRNA processing, pre-mRNA splicing factor. May have a role in pre-mRNA splicing; contains arginine/serine-rich domain and an RRM domain	0.04
#(NOCAT)		ENSP00000231844* Ecotropic virus integration 1 site protein.		0.04
X81789	Hs.77697:149	SF3A3 splicing factor 3a, subunit 3, 60kD	Nucleus, spliceosome, mRNA splicing, mRNA processing, pre-mRNA splicing factor. Spliceosome-associated protein 3a, subunit 3; component of the essential heterotrimeric splicing factor SF3a; contains a zinc finger	0.04

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
NM_002122	Hs.198253:21	HLA-DQA1 major histocompatibility complex, class II, DQ alpha 1	Pathogenesis, class II major histocompatibility complex antigen. Alpha 1 chain of HLA-DQ1 class II molecule (Ia antigen); complex binds peptides and presents them to CD4+ T lymphocytes[Proteome]	0.00
AB001914		Homo sapiens PACE4 gene, exon 23-25, complete cds	Function unknown	0.04
AA311919	Hs.69851:24	NOLA1 nuclear protein family A, member 1 (H1ACA small nuclear RNPs)	Involved in various aspects of rRNA processing and modification. Localize to the dense fibrillar components of nucleoli and to coiled (Cajal) bodies in the nucleus.	0.04
AI381750	Hs.283437:122,Hs.10 065:58	HTGN29 protein	Function unknown	0.04
#(NOCAT)		NM_000636*:Homo sapiens superoxide dismutase 2, mitochondrial (SOD2), mRNA, expression (RFX2), mRNA.	Mitochondrion, oxidative stress response, manganese superoxide dismutase. Manganese superoxide dismutase; Intramitochondrial free radical scavenging enzyme; has strong similarity to murine Sod2.	0.04
AA292998	Hs.163900:25	ESTs	Function unknown	0.04
BE439580	Hs.75498:40	SCYA20 small inducible cytokine subfamily A (Cys-Cys), member 20	Chemokine, chemotaxis, Immune response, signal transduction, extracellular space, cell-cell signalling, inflammatory response, antimicrobial humoral response. Cytokine A20 (exodus); chemotactic factor for lymphocytes, but not a chemotactic factor for monocytes	0.04
AI677897	Hs.76640:124	RGC32 RGC32 protein	Cytoplasm, cell cycle regulator, regulation of CDK activity. Strongly similar to RGC-32.	0.04
#(NOCAT)		Target Exon	Function unknown	0.04
N72403		Homo sapiens cDNA clone IMAGE:245132	Function unknown	0.05
BE003054	Hs.1695:46	MMP-12 matrix metalloproteinase 12 (macrophage elastase)	Zinc binding, cell motility, macrophage elastase, extracellular matrix, proteolysis and peptidolysis. Matrix metalloproteinase; degrades elastin	0.05
AL035588	Hs.153203:26,Hs.233 91:1	Human DNA sequence from clone 696P19 on chromosome 6p12.3-21.2. Contains the gene for TFEB, an NPM1 (Nucleophosmin, Numatrin) pseudogene and the MDF1 gene for MyoD family inhibitor (myogenic repressor 1-MF). Contains ESTs, STSs, GSSs and two putative CpG islands, complete sequence	Function unknown	0.05
AI080491	Hs.93270:3	ESTs, Moderately similar to S85657 alpha-1C- adrenergic receptor splice form 2 [H.sapiens]	Function unknown	0.05

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
AW770994	Hs.30340:125	hypothetical protein KIAA1165	Function unknown	0.05
H24177	Hs.75262:89,Hs.2389 12:1	CTSO cathepsin O	Cysteine-type endopeptidase, proteolysis and peptidolysis. Cathepsin O; cysteine (thiol) protease	0.05
AF149781	Hs.20450:29	BCM-like membrane protein precursor	Function unknown	0.05
NM_001955	Hs.2271:45,Hs.306:1	EDN1 endothelin 1	Circulation, peptide hormone, soluble fraction, signal transduction, extracellular space, cell-cell signalling, blood pressure regulation, positive control of cell proliferation, Preproendothelin 1; precursor of the hormone endothelin 1	0.05
A1680737	Hs.28908:204,Hs.32 6198:1	TCF4 transcription factor 4	Nucleus, RNA polymerase II transcription factor, transcription regulation from Pol II promoter. Transcriptional activator; interacts with ITF1 (TCF3); contains basic helix-loop-helix domain Proteome	0.05
A1752668	Hs.76869:183	NNMT nicotinamide N-methyltransferase	Nicotinamide N-methyltransferase; catalyzes the N-methylation of nicotinamide and other pyridines, structurally-related drugs and xenobiotics Proteome	0.05
AA505445	Hs.300697:21	IGHG3 immunoglobulin heavy constant gamma 3 (G3m marker)	Constant region of heavy chain of IgG3	0.05
BE246649; NM_003955	Hs.345728	SOC3 STAT induced STAT-inhibitor 3; suppressor of cytokine signalling 3	suppression of IL-6 mediated signalling	0.02
M86849	Hs.323733:62,Hs.300 816:5	GJB2 gap junction protein, beta 2, 28kD (connexin 26)	Hearing, connexon, plasma membrane, connexon channel, cell-cell signalling, small molecule transport. Connexin 26; gap junction protein expressed in various tissues including cochlea.	0.00
AW963419	Hs.155223:20	STC2 stanniocalcin 2	Peptide hormone, cell-cell signalling, glycopeptide hormone, nutritional response pathway, cell surface receptor linked signal transduction. Stanniocalcin 2; may regulate metal ion homeostasis and inhibits phosphate uptake.	0.00
BE298665	Hs.14846:132	Homo sapiens mRNA; cDNA DKFZp564D016 (from clone	Function unknown	0.00
AK000637	Hs.46624:11	HSPC043 HSPC043 protein	Function unknown	0.00
BE077546	Hs.31447:27	ESTs, Moderately similar to A46010 X-linked retinopathy protein [H.sapiens]	Function unknown	0.00
T97307		gb:ye53h05.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:121497 3,	Function unknown	0.00

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
		mRNA sequence.		
R24601	Hs.108300:46	Homo sapiens adenylosuccinate synthetase isozyme (ADSS) mRNA, complete cds	Function unknown	0.00
BE090176	Hs.179902:95	Interlin-CDW92 antigen	choline transporter-like protein	0.00
AA393907	Hs.97179:22	ESTs	Function unknown	0.00
W28729	Hs.236510:8	Homo sapiens mRNA; cDNA DKFZp666D074 (from clone DKFZp666D074)	Function unknown	0.00
BE313754	Hs.13350:52	Homo sapiens mRNA; cDNA DKFZp586D0918	Function unknown	0.01
AW673081	Hs.54828:9	ESTs	Function unknown	0.01
AA356694	Hs.94011:42, Hs.7744 :2, Hs.231043:1	HCA4 Hepatocellular carcinoma-associated protein HCA4	Function unknown	0.01
L08239	Hs.5326:11	MG61 Porcupine	amino acid system N transporter 2;	0.01
BE397849	Hs.94109:40	Homo sapiens cDNA FLJ34399 fis, clone HCHON2001359	Function unknown	0.01
NM_012317	Hs.45231:36	LDOC1 Leucine zipper, down-regulated in cancer 1	Nucleus, negative control of cell proliferation. Nuclear protein; contains a leucine zipper-like motif	0.01
NM_000947	Hs.74519:20	PRIM2A primase, polypeptide 2A (58kD)	DNA primase, DNA replication, priming, alpha DNA polymerase/primase complex. Subunit of DNA primase polypeptide 2A; part of the DNA polymerase alpha-primase complex	0.01
AJ250562	Hs.92749:133	Homo sapiens partial TM4SF2 gene for tetraspanin protein, exon 1 and joined CDS	Function unknown	0.01
AL040183	Hs.123484:24, Hs.326 906:1	Homo sapiens mRNA; cDNA DKFZp686E1934 (from clone DKFZp686E1934)	Function unknown	0.01
BE207573	Hs.83321:32	NIMB neuromedin B	Peptide hormone, soluble fraction, signal transduction, cell-cell signalling. Precursor of neuromedin B, a C-terminally amidated peptide hormone; similar to bombesin	0.01
BE564162	Hs.250820:45	FLJ14827 hypothetical protein FLJ14827	Function unknown	0.01



Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
BE439580	Hs.75498:40	SCYA20 Small inducible cytokine subfamily A (Cys-Cys), member 20	Chemokine, chemotaxis, immune response, signal transduction, extracellular space, cell-cell signalling, inflammatory response, antimicrobial humoral response. Cytokine A20 (exodus); chemotactic factor for lymphocytes, but not a chemotactic factor for monocytes	0.01
AW067800	Hs.155223:52	STC2 stanniocalcin 2	Peptide hormone, cell-cell signalling, glycopeptide hormone, nutritional response pathway, cell surface receptor linked signal transduction. Stanniocalcin 2; may regulate metal ion homeostasis and inhibits phosphate uptake.	0.01
AA569756	Hs.87803:10	Homo sapiens cDNA FLJ30158 fs, clone BRACE2000487	Function unknown	0.01
AW138180	Hs.180248:8	ZNF124 zinc finger protein 124 (HZF-16)	DNA binding. C2H2 zinc-finger protein 124	0.01
AF126245	Hs.14791:48	ACAD8 acyl-Coenzyme A dehydrogenase family, member 8	Lipid metabolism, acyl-CoA dehydrogenase. Member of the acyl-Coenzyme A dehydrogenase family; alpha,beta-dehydrogenates acyl-CoA esters	0.01
L10343	Hs.112341:98,Hs.1988:1	Homo sapiens elafin precursor, gene, complete cds	elastase-specific inhibitor in bronchial secretions	0.01
NM_002514	Hs.235935:38	NOV neuroblastoma overexpressed gene	Insulin-like growth factor receptor binding protein. Insulin-like growth factor binding protein; may play a role in nephrogenesis	0.01
AI663735	Hs.166755:3	ESTs	Function unknown	0.01
NM_005397	Hs.16428:160,Hs.248780:1	PODXL podocalyxin-like	Integral plasma membrane protein. Transmembrane protein similar to rodent podocalyxins	0.01
W26391	Hs.301206:100	KIF3B kinesin family member 3B	Plus-end kinesin, microtubule motor, anterograde axon cargo transport, plus-end-directed kinesin ATPase, determination of left-right asymmetry. Similar to murine Kif3b; may have a role in intracellular organelle transport, may act in left-right determination in embryogenesis; are a microtubule-associated motor protein	0.01
H15474	Hs.132898:166	FADS1 fatty acid desaturase 1	C-5 sterol desaturase, fatty acid desaturation. Integral membrane protein. Delta-5 desaturase; catalyzes production of polyenoic fatty acids such as arachidonic acid	0.01
U51168	Hs.173824:106	TDG Thymine-DNA glycosylase	DNA repair, nucleoplasm, damaged DNA binding, base-excision repair, GT-mismatch-specific thymine-DNA glycosylase. Thymine-DNA glycosylase; excises uracil and thymine from mispairs with guanine	0.01
AA243499	Hs.104800:23	FLJ10134 hypothetical protein FLJ10134	Highly similar to murine p19.5; are a membrane protein	0.01
AW408807	Hs.34497:46	FLJ22116 hypothetical protein FLJ22116	Function unknown	0.01

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
AI738719	Hs.198427:98	HK2 Hexokinase 2	Hexokinase, cell cycle control, glucose catabolism, glucose metabolism, mitochondrial outer membrane. Hexokinase II; converts aldo- and keto-hexose sugars to the hexose-6-phosphate	0.01
AB040888	Hs.41793:110	Homo sapiens mRNA for KIAA1455 protein, partial cds	Function unknown	0.01
BE313077	Hs.93135:40,Hs.228357:1	Homo sapiens cDNA FLJ39971 fs, clone SPLEN2028066	Function unknown	0.01
AI677897	Hs.76640:124	RGC32 RGC32 protein	Cytoplasm, cell cycle regulator, regulation of CDK activity. Strongly similar to RGC-32	0.01
C14898	Hs.192986:5	ESTs	Function unknown	0.01
AI821730	Hs.116524:7	Homo sapiens cDNA FLJ35800 fs, clone TEST12005933	Function unknown	0.01
AF007393	Hs.177574:111	PRKRIR protein-kinase, interferon-inducible double stranded RNA dependent inhibitor, repressor of (P58 repressor)	Stress response, protein binding, signal transduction, translational regulation, negative control of cell proliferation. Regulates interferon-induced protein kinase PKR (PRKR) activity by binding and inhibiting the PKR-regulator P58IPK (PRKR)	0.01
H65423	Hs.17631:42	DKFZP434E2135 hypothetical protein	Function unknown	0.01
N46243	Hs.110373:28	ESTs, Highly similar to T42828 secreted leucine-rich repeat-containing protein SLIT2 - mouse (fragment) [M.musculus]	Function unknown	0.01
AA095971	Hs.198793:56,Hs.309674:7	Homo sapiens cDNA: FLJ22463 fs, clone HRC10126	Function unknown	0.01
U20350	Hs.78913:33	CX3CR1 chemokine (C-X3-C) receptor 1	Virulence, chemotaxis, coreceptor, cell adhesion, plasma membrane, chemokine receptor, response to wounding, cellular defense response, integral plasma membrane protein, G-protein linked receptor protein signalling pathway. CX3C chemokine receptor, G protein-coupled receptor, mediates leukocyte migration and adhesion, binds the CX3C chemokine fractalkine and signals through a pertussis toxin sensitive G-protein	0.01
NM_005756	Hs.184942:18	GPR84 G protein-coupled receptor 64	Spermatogenesis, G-protein linked receptor, integral plasma membrane protein, G-protein linked receptor protein signalling pathway. Member of the G protein-coupled receptor family	0.01
D19589	Hs.13463:87	FLJ14753 hypothetical protein FLJ14753	Function unknown	0.02
AW957448	Hs.301711:74	ESTs	Function unknown	0.02

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
AW294647	Hs.233634:40	C20orf39 chromosome 20 open reading frame 39	Function unknown	0.02
BE159718	Hs.85335:46	Homo sapiens, clone IMAGE:4513159, mRNA	Function unknown	0.02
AI888490	Hs.55902:22	EDG3 endothelial differentiation, sphingolipid G-protein-coupled receptor, 3	Lipid binding, plasma membrane, inflammatory response, G-protein linked receptor, embryogenesis and morphogenesis, integral plasma membrane protein, positive control of cell proliferation, cytosolic calcium ion concentration elevation, G-protein linked receptor protein signalling pathway. Lysophospholipid receptor, a G protein-coupled receptor; activates calcium flux and serum response element driven transcription	0.02
AA022569	Hs.29802:35, Hs.271785:1	ESTs	Function unknown	0.02
BE147740	Hs.104558:21	ESTs, Moderately similar to hypothetical protein FLJ20378 [Homo sapiens]	Function unknown	0.02
AI798863	Hs.87191:8	ESTs	Function unknown	0.02
BE484341	Hs.21201:18	Interim-DKFP58650846: nectin 3	Low similarity to PVRL1; are a membrane glycoprotein; contains an immunoglobulin (Ig) domain	0.02
AL080235	Hs.35861:34, Hs.289068:1	RIS1 Ras-Induced senescence 1	Rat brain specific binding protein	0.02
AI557212	Hs.17132:102, Hs.330782:1	ESTs	Function unknown	0.02
X75208	Hs.2913:41	EPH3 EphB3	Signal transduction, integral plasma membrane protein, transmembrane receptor protein tyrosine kinase. Eph-related receptor tyrosine kinase B3	0.02
AA628980	Hs.192371:3	DSCR8 Down syndrome critical region protein DSCR8	Melanoma-testis-associated protein 2	0.02
BE242587	Hs.118651:39	HHEX hematopoietically expressed homeobox	Nucleus, DNA binding, transcription factor, developmental processes, antimicrobial humoral response. Member of the homeodomain family of DNA binding proteins; may regulate gene expression, morphogenesis, and differentiation	0.02
NM_005512	Hs.151841:65	GARP glycoprotein A repetitions predominant	Integral plasma membrane protein. Putative transmembrane cell surface protein; has an extracellular domain comprised largely of leucine-rich repeats	0.02
AW853853	Hs.292833:19	PAEP progesteragen-associated endometrial protein (placental protein 14, pregnancy-associated endometrial alpha-2-globulin, alpha uterine	Developmental processes. Placental protein 14 (Glycodein); member of lipocalin superfamily, highly similar to beta-lactoglobulins	0.02

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
AU076811	Hs.154672:122	MTMFD2 methylene tetrahydrofolate dehydrogenase (NAD dependent), methylenetetrahydrofolate cyclohydrolase	Mitochondrion, electron transporter, methylenetetrahydrofolate cyclohydrolase, methylenetetrahydrofolate dehydrogenase. NAD-dependent methylene tetrahydrofolate dehydrogenase-cyclohydrolase; may provide formyltetrahydrofolate for formylmethionyl tRNA synthesis; involved in initiation of mitochondrial protein synthesis	0.02
AW988613	Hs.79428:166	BNIP3 BCL2/adenovirus E1B 19kD-interacting protein 3	Anti-apoptosis, apoptosis inhibitor. Bcl2-related protein 3; binds antiapoptotic viral E1B 19 kDa protein and cellular Bcl2 protein	0.02
AL353944	Hs.50115:14	Homo sapiens mRNA; cDNA DKFZp761J1112 (from clone DKFZp761J1112)	Function unknown	0.02
BE614149	Hs.20814:29,Hs.306626:27	LOC51072: C21orf19-like protein	Function unknown	0.02
AA292998	Hs.163900:25	ESTs	Highly similar to winged helix/forkhead transcription factor	0.02
H12912	Hs.274691:138	AK3 adenylate kinase 3	Nucleobase, nucleoside, nucleotide and nucleic acid metabolism. Adenylate kinase 3; strongly similar to murine Ak4	0.02
AA188763	Hs.36793:4	SLC12A8 solute carrier family 12 (potassium/chloride transporters), member 8	Solute carrier family 12 (potassium/chloride transporters), member 8	0.02
AK000598	Hs.3818:56	HPCAL1 hippocalcin-like 1	Calcium-binding protein with similarity to hippocalin (human HPCA); expressed only in the brain.	0.02
AI970797	Hs.64859:16	ESTs	Function unknown	0.02
AW519204	Hs.40808:22	ESTs	Function unknown	0.02
Z42387	Hs.83883:114	TMEPA1 transmembrane, prostate androgen induced RNA	Function unknown	0.02
AF145713	Hs.61480:51	SCHIP1 schwannomin-interacting protein 1	Cytoplasm. Associates with the neurofibromatosis type 2 protein schwannomin (NF2); contains a coiled-coil domain	0.02
AA972412	Hs.13755:41	FBXW2 f-box and WD-40 domain protein 2	Protein modification, ubiquitin-protein ligase, proteolysis and peptidolysis, ubiquitin conjugating enzyme. F-box and WD-40 domain protein 2; putative SCF ubiquitin ligase subunit involved in protein degradation; contains a WD-40 domain and an F-box	0.02
AK001564	Hs.104222:139,Hs.296267:4	Homo sapiens cDNA FLJ10702 fis, clone NT2RP3000759, weakly similar to ADP-RIBOSYLATION FACTOR	Member of the ADP-ribosylation factor (ARF) family; putative GTP-binding protein involved in protein trafficking	0.02

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
AW959861	Hs.290943:28	ESTs	Function unknown	0.02
BE313555	Hs.7252:158	RAI17 retinoic acid induced 17	Function unknown	0.02
W25005	Hs.24395:189	zb67e02.r1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone IMAGE:308666 5', mRNA sequence ESTs	Function unknown	0.02
AI193356	Hs.160316:3		Function unknown	0.02
AF111108	Hs.3382:223	PPP4R1 Protein phosphatase 4, regulatory subunit 1	Protein phosphatase	0.02
AI130740	Hs.6241:116	PIK3R1 phosphoinositide-3-kinase, regulatory subunit, polypeptide 1 (p85 alpha)	A family of enzymes that phosphorylate the 3'-hydroxyl of phosphatidylinositol (PtdIns).	0.02
AA985190	Hs.246875:42	FLJ20059 hypothetical protein FLJ20059	Contains four Kelch motif domains	0.02
BE221880	Hs.288555:144	XRN2 5'-3' exonuclease 2	Nucleus, nuclease, recombination, RNA catabolism, RNA processing. 5'-3' Exoribonuclease; similar to Schizosaccharomyces pombe Dhp1p	0.03
AF084545		Homo sapiens versican Vint isoform, mRNA, partial cds	Function unknown	0.03
R26594	Hs.287893:43	TAPBP-R: TAP binding protein related	Has low similarity to TAPBP (Tapasin); contains two immunoglobulin (Ig) domains Proteome	0.03
AW247380	Hs.12124:116	ELAC2 elac homolog 2 (E. coli)	putative prostate cancer susceptibility protein	0.03
AA384261	Hs.131365:7	ESTs	Weakly similar to T31613 hypothetical protein Y50E8A.J - Caenorhabditis elegans [C.elegans]	0.03
U25849	Hs.75393:141	ACP1 Human red cell-type low molecular weight acid phosphatase (ACP1) gene, exon 6 and 7, complete cds	Acid phosphatase	0.03
AF282892	Hs.123159:14	SPAG4 Sperm associated antigen 4	Spermatogenesis, structural protein. Sperm associated antigen 4; predicted ortholog of rat SPAG4, which interacts with rat ODF27, the 27kDa outer dense fiber protein of elongating spermatids	0.03
AW342140	Hs.182545:1	ESTs, Weakly similar to POL2_MOUSE Retrovirus-related POL polypeptide	Function unknown	0.03
AL133572	Hs.189009:58	PCCX2 protein containing CXXC domain 2	DNA-binding protein with PHD finger and CXXC domain, is regulated by proteolysis.	0.03

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
AI497778	Hs.20509:4	HBXAP Hepatitis B virus x associated protein	Weakly similar to Drosophila CG8677	0.03
AI745379	Hs.42811:31	TAF13 TAF13 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 18 kD	TFIID complex, protein binding, transcription factor, general RNA polymerase II transcription factor. TBP-associated factor K; component of TFIID complexes containing TAF130 (TAF2H). The protein encoded by this gene is a lung cancer associated protein. The function of the protein is not known. Multiple alternatively spliced transcript variants have been described for this gene but some of their full length sequence has not been determined.	0.03
U51712	Hs.13776:135	LAGY: lung cancer-associated Y protein	Function unknown	0.03
AW375974	Hs.156704:4	ESTs	GAGE genes are expressed in a variety of tumors and in some fetal and reproductive tissues. This gene is strongly expressed in Ewing's sarcoma, alveolar rhabdomyosarcoma and normal testis. The protein encoded by this gene contains a nuclear localization signal and shares a sequence similarity with other GAGE/PAGE proteins. Because of the expression pattern and the sequence similarity, this protein also belongs to a family of CT (cancer-testis) antigens.	0.03
AF251237	Hs.112208:16	GAGED2 G antigen, family D, 2	Mitochondrion, oxidative stress response, manganese superoxide dismutase, Manganese superoxide dismutase; intramitochondrial free radical scavenging enzyme; has strong similarity to murine Sod2.	0.02
NM_000638		Homo sapiens superoxide dismutase 2, mitochondrial (SOD2), mRNA, expression (RFX2), mRNA.	Function unknown	0.01
AA130986	Hs.271627:1	ESTs	Function unknown	0.01
AA218363	Hs.262858:48,Hs.327737:2	DKFZP434B044 hypothetical protein DKFZp434B044	Function unknown	0.00
AA628980	Hs.192371:3	DSCR8 down syndrome critical region protein DSCR8	Function unknown	0.02
AA811857	Hs.220913:9	Homo sapiens cDNA FLJ40827 fis, clone TRACH2011600	Function unknown	0.01
AA897108		gbam08a08.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone 3', mRNA sequence	Function unknown	0.02
AB040888	Hs.41793:110	Homo sapiens mRNA for KIAA1455 protein, partial cds	Function unknown	0.01

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
AF212225	Hs.283693:104	Homo sapiens BM022 mRNA, complete cds	Function unknown	0.02
A089575	Hs.9071:52	ESTs	Function unknown	0.02
A1282028	Hs.25205:10	ESTs	Function unknown	0.02
A1388826	Hs.30654:15	FLJ10849: hypothetical protein FLJ10849	Moderately similar to members of the septin family	0.02
A1718702	Hs.308026:11,Hs.194490:8	HLA-DRB3 major histocompatibility complex, class II, DR beta 5	Signal transduction, integral plasma membrane protein, class II major histocompatibility complex antigen. Beta 3 chain of HLA-DR; subunit of MHC class II molecule, complex binds peptides and presents them to CD4+ T lymphocytes	0.02
A1827248	Hs.224398:3	Homo sapiens cDNA FLJ11469 fis, clone HEMBA1001658	Function unknown	0.01
AK002039	Hs.26243:38	MRV1 murine retrovirus integration site 1 homolog	Oncogenesis, tumor suppressor, endoplasmic reticulum membrane. Similar to human MLRP, may act as a tumor suppressor	0.02
AL109791	Hs.241559:3	Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 151432	Function unknown	0.00
AW090198	Hs.4779:29	LOC127829: hypothetical protein BC015408	Function unknown	0.01
AW296454	Hs.24743:92	FLJ20171: hypothetical protein FLJ2017	Contains three RNA recognition motifs (RRM, RBD, or RNP)	0.02
AW445034	Hs.256578:4	ESTs	Function unknown	0.00
AW452948	Hs.257631:3	ESTs	Function unknown	0.01
AW470411	Hs.288433:27	HNT: neurotrophin	Cell adhesion, neuronal cell recognition, integral plasma membrane protein. Neurotrophin; may function as a GPI-anchored neural cell adhesion molecule; member of the immunoglobulin superfamily	0.02
AW885727	Hs.301570:22	FST follistatin	Developmental processes. Follistatin; inhibits the release of follicle-stimulating hormone (FSH)	0.01
AW970859	Hs.313503:4	ESTs	Function unknown	0.02
AW979189	Hs.283367:3	ESTs	Function unknown	0.01

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
BE165866	Hs.83623:66	Human XIST, coding sequence "a" mRNA (locus DXS399E)	XIST mRNA	0.01
BE175582		gb:RC5-HT0580-100500-022-C01 HT0580 Homo sapiens cDNA, mRNA sequence	Function unknown	0.01
BE242587	Hs.118651:39	HHEX hematopoietically expressed homeobox	Nucleus, DNA binding, transcription factor, developmental processes, antimicrobial humoral response. Member of the homeodomain family of DNA binding proteins; may regulate gene expression, morphogenesis, and differentiation	0.01
BE271927	Hs.87385:31,Hs.3079 40:4	LOC115416: hypothetical protein BC012331	Function unknown	0.01
BE439580	Hs.75498:40	SCYA20 small inducible cytokine subfamily A (Cys-Cys), member 20	Chemokine, chemotaxis, immune response, signal transduction, extracellular space, cell-cell signalling, inflammatory response, antimicrobial humoral response. Cytokine A20 (exodus); chemotactic factor for lymphocytes, but not a chemotactic factor for monocytes	0.02
BE464016	Hs.238856:35	Homo sapiens cDNA FLJ37793 fis, clone BRHIP3000473	Function unknown	0.02
D63218	Hs.153684:137	FRZB frizzled-related protein	Membrane, extracellular, skeletal development. Frizzled-related protein; similar to frizzled family of receptors	0.02
F34856	Hs.292457:120	Homo sapiens, clone MGC:16362 IMAGE:3927795, mRNA, complete cds	Function unknown	0.02
M83822	Hs.62354:112	LRBA LPS-responsive vesicle trafficking, beach and anchor containing	May mediate protein-protein interactions; contains two WD domains (WD-40 repeats) and a beige/BEACH domain	0.02
N33937	Hs.10336:6	ESTs	Function unknown	0.01
N49088	Hs.93986:4	ESTs	Function unknown	0.01
N51357	Hs.260855:62	NSE1: NSE1	Function unknown	0.02
N80486	Hs.39911:17	Homo sapiens mRNA for FLJ00089 protein, partial cds	Function unknown	0.02
NM_000954	Hs.8272:265,Hs.3323 55:1	PTGDS prostaglandin D2 synthase (21kD, brain)	Membrane, prostaglandin-D synthase. Glutathione-independent prostaglandin D2 synthase; membrane associated, catalyzes synthesis of prostaglandin D; member of the lipocalin family of transporters	0.02
NM_005756	Hs.184942:18	GPR64 G protein-coupled receptor 64	Spermatogenesis, G-protein linked receptor, integral plasma membrane protein, G-protein linked receptor protein signalling pathway. Member of the G protein-coupled receptor family	0.02



Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
NM_016652	Hs.268281:61	CRNKL1 Cm, crooked neck-like 1 (Drosophila)	Function unknown	0.02
R26584	Hs.267993:43	TAPBP-R; TAP binding protein related	Has low similarity to TAPBP (Tapasin); contains two immunoglobulin (Ig) domains	0.01
R31178	Hs.287820:8	FN1 fibronectin 1	Cell adhesion, cell motility, cell adhesion, soluble fraction, signal transduction, extracellular matrix, extracellular space. Fibronectin 1; member of family of proteins found in plasma and extracellular matrix	0.02
W05391	Hs.83623:8	Homo sapiens cDNA FLJ30298 fis, clone BRACE2003172	Function unknown	0.02
W25005	Hs.24395:199	zb87e02.r1 Soares fetal_lung_NbHL19W Homo sapiens cDNA clone IMAGE:308668 5' mRNA sequence	Function unknown	0.01
W45393	Hs.55888:15	ATP7 activating transcription factor 7	Transcription factor. Leucine zipper DNA-binding protein; recognizes a cAMP response element (CRE), involved in the regulation of adenovirus Ela-responsive and cellular cAMP-inducible promoters	0.02
W68815	Hs.301885:20	Homo sapiens cDNA FLJ33794 fis, clone CTONG1000008	Function unknown	0.01
X65965		H. sapiens SOD-2 gene for manganese superoxide dismutase	Mitochondrion, oxidative stress response, manganese superoxide dismutase. Manganese superoxide dismutase; intramitochondrial free radical scavenging enzyme; has strong similarity to murine Sod2.	0.01
X76732	Hs.3164:58	NUCB2 nucleobindin 2	Cytosol, DNA binding, plasma membrane, calcium binding, extracellular space. Nucleobindin 2; may bind DNA and calcium; has DNA-binding and EF-hand domains, and a leucine-zipper	0.02
Z45051	Hs.22920:25	C20orf103 chromosome 20 open reading frame 103	Low similarity to a region of murine Lamp1/Proteome	0.02

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	P value
<b>c. downregulated genes</b>				
NM_022117	Hs.138164;23	SE20-4, cutaneous T-cell lymphoma-associated tumor antigen se20-4se20-4	Cutaneous T-cell lymphoma-associated tumor antigen se20-4se20-4; differentially expressed nuclear TGF-beta1 target protein (DENTT); also known as CDA1	0
NM_005460	Hs.24948;32, Hs.300445;4	SNCAIP, synuclein, alpha interacting protein (synphilin)	Cytoplasm, pathogenesis, protein binding. Synphilin-1; promotes formation of cytosolic inclusions in neurons (SNCAIP). Synuclein alpha interacting protein contains several protein-protein interaction domains and interacts with alpha synuclein in neurons. Mutations of SNCAIP have been linked to Parkinson disease.	0
NM_002387	Hs.1345;5	MCC, mutated in colorectal cancers	Receptor, signal transduction, tumor suppressor. Similar to the G protein-coupled m3 muscarinic acetylcholine receptor. MCC is a candidate for the putative colorectal tumor suppressor gene. The MCC gene product are involved in early stages of colorectal neoplasia in both sporadic and familial tumors.	0
A1745249	Hs.23650;30	Homo sapiens, clone MGC:9888 IMAGE:3868330	Function unknown	0.0009
A1894200	Hs.356620, Hs.227913;11	ESTs	Function unknown	0.0442

Table 2  
Genes having modified expression in serous ovarian cancer relative to normal ovarian tissue

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
M25809	Hs.84173	ATP6V1B1, ATPase, H <sup>+</sup> transporting, lysosomal 59/68kD, V1 subunit B, isoform 1 (Renal tubular acidosis with deafness)	Subunit B1 (beta subunit) of a vacuolar-type H <sup>+</sup> -ATPase 1; apical proton pump that mediates distal nephron acid secretion	1082.30
AW959311	Hs.172012	DKFZP434J037: hypothetical protein DKFZP434J037	Function unknown	227.83
H16423	Hs.82685	Homo sapiens mRNA; cDNA DKFZP313F0317 (from clone DKFZP313F0317)	Function unknown	74.54
A1733848	Hs.71835	ZNF338, zinc finger protein 338	Zinc finger protein	55.13
AW055308	Hs.31803	NAC1, transcriptional repressor NAC1	Function unknown	52.63
AF034102	Hs.32951	SLC29A2, solute carrier family 29 (nucleoside transporters), member 2	Nitrobenzylthioinosine-Inosensitive equilibrative nucleoside transporter 2; may act in the uptake of purine and pyrimidine nucleosides	44.34
A1791805	Hs.85549	FLJ20273: RNA-binding protein	Contains four RNA recognition motifs (RRM, RBD, or RNP)	43.21
AW286454	Hs.24743	FLJ20171: hypothetical protein FLJ20171	Contains three RNA recognition motifs (RRM, RBD, or RNP)	38.91
Z43989	Hs.82141	Human clone 23612 mRNA sequence	Function unknown	37.89
AL043980	Hs.7886	PEL11, pellino homolog 1 (Drosophila)	Pellino protein	35.20
BE514982	Hs.38991	S100A2, S100 calcium binding protein A2	S100 calcium-binding protein A2; interacts with target proteins to link extracellular stimuli and cellular responses; member of the S100 tissue/cell specific Ca <sup>2+</sup> -binding protein family	34.53
AI811807	Hs.108649	Target Exon Homo sapiens cDNA FLJ12534 fs, clone NT2RM400244	Function unknown	34.02
U90441	Hs.3622	P4HA2, procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha polypeptide II	Function unknown	32.34
T98226	Hs.171952	OCLN, occludin	Alpha 2 subunit of prolyl 4-hydroxylase; catalyzes the formation of 4-hydroxyproline in collagens	32.24
			This gene encodes an integral membrane protein which is located at tight junctions. This protein are involved in the formation and maintenance of the tight junction.	31.56

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
R35343	Hs.24968	Human DNA sequence from clone RP1-233G18 on chromosome Xq22.1-23. Contains the 6' part of a novel gene, ESTs, STSs, GSSs and a putative CpG island		31.22
BE247295	Hs.78452	SLC20A1, solute carrier family 20 (phosphate transporter), member 1	Sodium-dependent phosphate symporter; acts as a cell-surface receptor for gibbon ape leukemia virus	30.16
AB037734	Hs.4983	PCDH18, protocadherin	Protocadherin	29.90
		C6000394*:g 12737280 ref XP_006682.2  keratin 18 [Homo sapiens]  6833	Function unknown	29.30
AF212223	Hs.25010	Homo sapiens BM025 mRNA, complete cds	Function unknown	28.85
AA902656	Hs.21943	NIF3L1, NIF3 (Ngg1 interacting factor 3, S.pombe homolog)-like 1	Anyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 1	27.73
X14008	Hs.234734	Human lysozyme gene (EC 3.2.1.17)	Lysozyme	27.66
AA570256		LOC116238: hypothetical protein BC014072	Function unknown	27.52
AA137152	Hs.286048	PSA, phosphoserine aminotransferase	The protein encoded by this gene is likely a phosphoserine aminotransferase, based on similarity to proteins in mouse, rabbit, and Drosophila. Alternative splicing of this gene results in two transcript variants encoding different isoforms.	25.57
BE621807		TM4SF1, transmembrane 4 superfamily member 1	L6 antigen; member of the transmembrane 4 superfamily (TM4SF)	25.40
AB041038	Hs.57771	KLK11, kallikrein 11	Trypsin-like serine protease; has serine protease activity	25.05
F13386	Hs.7888	Homo sapiens clone 23738 mRNA sequence	Function unknown	22.50
AA158177	Hs.118722	FUT8, fucosyltransferase 8 (alpha (1,6) fucosyltransferase)	N-linked glycosylation, oligosaccharide biosynthesis, glycoprotein 6-alpha-L-fucosyltransferase, Alpha(1,6)fucosyltransferase (GDP-L-Fuc:N-acetyl-beta-D-glucosaminide:alpha(1-6) fucosyltransferase); transfers fucose to N-linked type complex glycopeptides from GDP-Fuc; functions in asparagine-linked glycoprotein oligosaccharide synthesis	21.90
BE267045	Hs.75084	TBCC, tubulin-specific chaperone c	Tubulin-specific chaperone c; cofactor in the folding pathway of beta-tubulin, mediates the release of beta-tubulin polypeptides committed to the native state	21.49
		NM_005936:Homo sapiens myeloid/lymphoid or mixed-lineage leukemia (trithorax (Drosophila) homolog); translocated to, 4 (MLLT4), mRNA.	Function unknown	20.46

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
AA150864	Hs.780	MGST1, microsomal glutathione S-transferase 1	Microsome, glutathione transferase. Microsomal glutathione S-transferase; catalyzes the conjugation of glutathione to electrophilic compounds; member of a family of detoxication enzymes.	20.35
AW855632	Hs.66686	EST367702. MAGE resequences, MAGD Homo sapiens cDNA, mRNA sequence	Function unknown	20.26
AW837046	Hs.6527	QV1-LT0037-150200-089-e09 LT0037 Homo sapiens cDNA, mRNA sequence	Function unknown	19.60
AA286987	Hs.24724	MFHAS1, malignant fibrous histiocytoma amplified sequence 1	The primary structure of its product includes an ATP/GTP-binding site, three leucine zipper domains, and a leucine-rich tandem repeat, which are structural or functional elements for interactions among proteins related to the cell cycle, and which suggest that overexpression might be oncogenic with respect to MFH.	19.16
AW401864	Hs.18720	PDCD8, programmed cell death 8 (apoptosis-inducing factor)	Mitochondrial apoptosis-inducing factor; flavoprotein inducing chromatin condensation and DNA fragmentation	19.01
AA186241	Hs.73980	zp9803.r1 Stragene muscle 937208 Homo sapiens cDNA clone IMAGE:628253 5' similar to gb:U18308 TROPONIN T, SLOW SKELETAL MUSCLE ISOFORMS (HUMAN); mRNA sequence	Function unknown	18.82
NM_004998	Hs.82251	MYO1E, myosin IE	Highly similar to class I myosin; may bind proline-rich peptides; contains an Src homology 3 (SH3) and myosin head domain (motor domain)	18.62
AW873704	Hs.320831	C20orf72: chromosome 20 open reading frame 72	Function unknown	18.19
AW361668	Hs.49500	KIAA0746: KIAA0746 protein	Function unknown	18.05
BE174595	Hs.368	PTS, 6-pyruvoyltetrahydropterin synthase	6-Pyruvoyltetrahydropterin synthase; synthesizes tetrahydrobiopterin, activity requires sepiapterin reductase, Mg <sup>2+</sup> , and NADPH	17.28
M31669	Hs.1735	Human inhibin beta-B-subunit gene, exon 2, and complete cds	Function unknown	16.24
AK001714	Hs.95744	FLJ10852, hypothetical protein similar to ankyrin repeat-containing protein AKR1	Are involved in protein-protein interactions; has five ankyrin repeats and a DHC-type zinc finger or NEW1 domain	16.09
AU076517	Hs.184276	AU076517 Sugano cDNA library Homo sapiens cDNA clone ClF3365 similar to 5'-end region of Homo sapiens ezrin-radixin-moesin binding phosphoprotein-60 mRNA, mRNA sequence	Function unknown	16.05

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
NM_006486	Hs.288215	STHM, sialyltransferase	Low similarity to beta-galactosidase a-2,3-sialyltransferase SIAT4B; member of the sialyltransferase family	15.93
BE148235	Hs.183063	Homo sapiens cDNA FLJ14201 fis, clone NT2RP3002866	Function unknown	15.91
AV653729	Hs.8185	SQDL: sulfide dehydrogenase like (yeast)	Sulfide dehydrogenase like	15.35
AL119871	Hs.1420	FGFR3, fibroblast growth factor receptor 3 (achondroplasia, thanatophoric dwarfism)	Fibroblast growth factor receptor 3; receptor tyrosine kinase that binds acidic and basic FGF	14.62
AA393071	Hs.182579	LAP3, leucine aminopeptidase	Leucine aminopeptidase	14.60
AL048753	Hs.303649	CCL2, chemokine (C-C motif) ligand 2	Cytokine A 2; chemotactic factor for monocytes	14.37
AI868872	Hs.282804	CP, ceruloplasmin (ferroxidase)	Ceruloplasmin; ferrous oxidase, binds copper in plasma and maintains iron homeostasis	14.07
NM_004419	Hs.2128	DUSP5, dual specificity phosphatase 5	Mitogen inducible dual specificity protein phosphatase 5; dephosphorylates extracellular signal-regulated kinase	14.05
AW969587	Hs.86368	EST381684 MAGE resequences, MAGK Homo sapiens cDNA, mRNA sequence	Function unknown	13.75
AW161449	Hs.72280	WNT7A, wingless-type MMTV Integration site family, member 7A	Very strongly similar to murine Wnt7a; may have a role in limb development and sexual dimorphism; member of the Wnt family of cell signalling proteins	13.48
BE409838	Hs.194657	CDH1, cadherin 1, type 1, E-cadherin (epithelial)	E-cadherin (uvomorulin); Ca <sup>2+</sup> -dependent glycoprotein, mediates cell-cell interactions in epithelial cells	12.92
BE540274	Hs.239	FOXM1, forkhead box M1	Cell-cycle regulated HNF-3/fork head; a transcriptional regulator	12.86
AF022375	Hs.73793	VEGF, vascular endothelial growth factor	Vascular endothelial growth factor; induces endothelial cell proliferation and vascular permeability	12.79
AW369278	Hs.23412	FLJ20160: hypothetical protein FLJ20160	Function unknown	12.73
AF147204	Hs.89414	CXCR4, chemokine (C-X-C motif), receptor 4 (fusin)	CXC chemokine receptor (fusin); G protein-coupled receptor binds CXC cytokines, mediates intracellular calcium flux	12.58
BE242818	Hs.311609	DDX39, DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 39	Strongly similar to human D6S91E; member of the DEAD/H box ATP-dependent RNA helicase family	12.43
NM_014791	Hs.184339	MELK, maternal embryonic leucine zipper kinase	Leucine zipper kinase	12.25
U38847	Hs.151518	TARBP1, TAR (HIV) RNA binding protein 1	Binds to the HIV-1 TAR RNA regulatory element, may function alone or with HIV-1 Tat to disengage RNA polymerase II during transcriptional elongation; has a leucine zipper	12.22
AW953575	Hs.303125	EST365945 MAGE resequences, MAGC Homo sapiens cDNA, mRNA sequence	Function unknown	12.21

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
A1949095	Hs.67776	ESTs, Weakly similar to T22341 hypothetical protein F47B8.5 - <i>Caenorhabditis elegans</i> [C.elegans]	Homo sapiens, clone IMAGE:5455669, mRNA, partial cds	12.08
BE274530	Hs.273333	FLJ10986, hypothetical protein FLJ10986	Member of the FGGY carbohydrate kinase family	11.75
AB020676	Hs.21543	KIA00899 protein Target Exon	Function unknown	11.73
H48289	Hs.26126:33	CLDN10, claudin 10	Function unknown	11.69
T34530	Hs.4210	Homo sapiens cDNA FLJ13069 fis, clone NT2RP3001752	Cell adhesion, integral plasma membrane protein, tight junction.	11.67
NM_022454	Hs.97984	SOX17, SRY (sex determining region Y)-box 17	Function unknown	11.50
AA737033	Hs.7165	Homo sapiens, clone IMAGE:4428577, mRNA, partial cds	SRY-related HMG-box transcription factor SOX17	11.42
AA433988	Hs.98502:8	MUC18, mucin 18, CA125	Function unknown	10.79
H91282	Hs.286232	Homo sapiens cDNA: FLJ23190 fis, clone LNG12190	Mucin 18, Alias CA125 ovarian cancer antigen	10.52
AW005054	Hs.47883	LOC57118: CamK1-like protein kinase	Function unknown	10.50
X69699	Hs.73149	PAX8, paired box gene 8	CamK1-like protein kinase; granulocyte-specific protein kinase that activates ERK/MAP kinase activity; similar to Ca(2+)-calmodulin-dependent kinase I (CamK1)	10.49
AW382987	Hs.86474:42	Homo sapiens cDNA, mRNA sequence	Member of the paired domain family of nuclear transcription factors; are involved in the ribosome assembly, required for normal thyroid development	10.39
AW957446	Hs.301711	Homo sapiens, clone MGC:23936 IMAGE:3838595, mRNA, complete cds	Function unknown	10.21
AA361562	Hs.178761	POH1: 26S proteasome-associated pad1 homolog	Function unknown	10.12
AA834626		RAD54L, RAD54 ( <i>S.cerevisiae</i> )-like	Ubiquitin-dependent protein degradation	10.01
A1878927	Hs.79284	MEST, mesoderm specific transcript (mouse) homolog	Has likely roles in mitotic and meiotic DNA recombination and repair; member of SNF2/SWI2 family of DNA-dependent ATPases	9.85
AW074266	Hs.23071	LOC85438: stonin 2	Mesoderm specific protein; member of the alpha/beta hydrolase fold family	9.83
NM_000947	Hs.74519	PRIM2A, primase, polypeptide 2A (58KD)	Stonin 2 Subunit of DNA primase polypeptide 2A; part of the DNA polymerase alpha-primase complex	9.74 9.72

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
NM_009187	Hs.56009	OAS3, 2'-5'-oligoadenylate synthetase 3 (100 kD)	Member of the 2'-5'-oligoadenylate synthetase family	9.68
AW276858	Hs.81256	S100A4, S100 calcium binding protein A4 (calcium protein, calvasculin, metastasin, murine placental homolog)	Calcylin (metastasis-associated) (S100 calcium-binding protein A4); interacts with targets to link extracellular stimuli and cellular responses; member of the S100 family of tissue-specific calcium-binding proteins	9.66
T18997	Hs.180372	LOC139231: hypothetical protein BC016683	Function unknown	9.49
AA262284	Hs.180383	DUSP6, dual specificity phosphatase 6	Dual specificity protein phosphatase 6; selectively dephosphorylates and inactivates MAP kinase	9.48
AA220238	Hs.94986	RPP38: ribonuclease P (38kD)	Nucleus, ribonuclease P. Subunit p38 of ribonuclease P	9.41
AW505308	Hs.75812	PCK2, phosphoenolpyruvate carboxykinase 2 (mitochondrial)	Phosphoenolpyruvate carboxykinase 2; forms phosphoenolpyruvate by decarboxylation of oxaloacetate at the rate-limiting step of gluconeogenesis	9.38
AI186431	Hs.288638	PLAB: prostate differentiation factor	Macrophage inhibitory cytokine, member of a subgroup of the TGF-beta superfamily	9.12
AI095718	Hs.135015	Homo sapiens cDNA FLJ40908 fis, clone UTERU2004698, highly similar to Mus musculus mRNA for thrombospondin type 1 domain	Function unknown	9.04
W70171	Hs.75938	UMPK, uridine monophosphate kinase	The protein encoded by this gene catalyzes the phosphorylation of uridine monophosphate to uridine diphosphate. This is the first step in the production of the pyrimidine nucleoside triphosphates required for RNA and DNA synthesis. In addition, an allele of this gene may play a role in mediating nonhumoral immunity to Hemophilus influenzae type B.	8.97
AI580935	Hs.105698	Homo sapiens cDNA FLJ31553 fis, clone NT2RI2001178	Function unknown	8.90
AB040914	Hs.278628	ShimL: Shroom-related protein	Shroom-related protein	8.87
AU076811	Hs.154672	MTFHD2, methylene tetrahydrofolate dehydrogenase (NAD+-dependent), methenyltetrahydrofolate cyclohydrolase	NAD-dependent methylene tetrahydrofolate dehydrogenase-cyclohydrolase; may provide formyltetrahydrofolate for formylmethionyl tRNA synthesis; involved in initiation of mitochondrial protein synthesis	8.71
AI089860	Hs.323401	LOC84661: dpy-30-like protein	dpy-30-like protein	8.71
D13688	Hs.136348:228, Hs.80988:2	OSF-2: osteoblast specific factor 2 (fasciclin I-like)	Cell adhesion, skeletal development. Putative bone adhesion protein; similar to the insect protein fasciclin I	8.64
AI788863	Hs.87191	ESTs	Function unknown	8.52



Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
U78093	Hs.15154	SRPX, sushi-repeat-containing protein, X chromosome	Putative membrane protein with short consensus repeat (sushi) domains	8.51
AI669780	Hs.188881	ESTs	Function unknown	8.37
AJ375726	Hs.279918	MGC2198: hypothetical protein MGC2198	Function unknown	8.37
AW271106	Hs.133294	ESTs	Function unknown	8.30
AK001782	Hs.15093	HSPC195: hypothetical protein HSPC195	Function unknown	8.18
AF019226	Hs.8036	RAB3D, RAB3D, member RAS oncogene family	GTP-binding protein; are involved in vesicle transport; member of the RAB family of small GTPases	7.94
AW868343	Hs.24255	LOC150898: prominin-related protein	Prominin-related protein	7.90
AF111858	Hs.105039	SLC34A2, solute carrier family 34 (sodium phosphate), member 2	Sodium-dependent phosphate transporter; member of the renal type II co-transporter family	7.87
AA863380	Hs.26040	Homo sapiens, clone MGC:40051 IMAGE:5243005, mRNA, complete cds	Function unknown	7.75
NIM_005764	Hs.271473	DP98: epithelial protein up-regulated in carcinoma, membrane associated protein 17	Up-regulated in malignant epithelial cells of renal cell carcinomas, and in carcinomas of colon, breast and lung	7.75
AW360901	Hs.183047	MGC4399: mitochondrial carrier protein	Mitochondrial carrier protein MGC4399	7.71
AL353944	Hs.50115	Homo sapiens mRNA; cDNA DKFZp761J1112 (from clone DKFZp761J1112)	Function unknown	7.69
H59789	Hs.42644	TXNL2, thioredoxin-like 2	Member of the thioredoxin family; has region of moderate similarity to glutaredoxin-like proteins	7.65
NIM_002984	Hs.75703	CCL4, chemokine (C-C motif) ligand 4	Cytokine A4	7.64
AA642452	Hs.130881	BCL11A, B-cell CLL/lymphoma 11A (zinc finger protein)	May bind nucleic acids; contains three C2H2 type zinc finger domains	7.61
AA789081	Hs.4029	GAS41: glioma-amplified sequence-41	Similar to the transcription factors AF-9 and ENL	7.46
H13032	Hs.103378	MGC11034, hypothetical protein MGC11034	Function unknown	7.42
BE384836	Hs.3454	KIAA1821: KIAA1821 protein	KIAA1821 protein	7.40
AW067800	Hs.155223	STC2, stanniocalcin 2	Stanniocalcin 2; may regulate metal ion homeostasis and inhibits phosphate uptake	7.38
T55979	Hs.116474	RFC3, replication factor C (activator 1) 3 (38kD)	Subunit of replication factor C (activator 1) 3; activator of DNA polymerases	7.35
AJ278018	Hs.55565	ANKRD3, ankyrin repeat domain 3	Ortholog of mouse protein kinase C-associated kinase, putative gene, ankyrin like, possible dual-specificity Ser/Thr/Tyr kinase domain	7.25

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
		NM_025080: Homo sapiens hypothetical protein FLJ22316 (FLJ22316), mRNA, VERSION NM_025079.1 GI:13376631	Function unknown	7.22
AA084248	Hs.65939.64	GPR39, G protein-coupled receptor 39	GPR39, G protein-coupled receptor 39	7.15
BE620738	Hs.173125	PP1F, peptidylprolyl isomerase F (cyclophilin F)	Cyclophilin F (peptidylprolyl isomerase F); binds the immunosuppressant drug cyclosporin A	7.06
AF072873	Hs.114218	FZD8, frizzled (Drosophila) homolog 8	Frizzled-8; may function in tissue polarity, development and carcinogenesis; similar to frizzled receptor family, has seven transmembrane domains	7.04
AA852773	Hs.334838	KIAA1866 protein	KIAA1866 protein	6.99
R07566	Hs.73817	CCL3, chemokine (C-C motif) ligand 3	Macrophage inflammatory protein 1 alpha; chemokine	6.98
NM_005211	Hs.174142	CSF1R, colony stimulating factor 1 receptor, formerly McDonough feline sarcoma viral (v-fms) oncogene homolog	Macrophage colony stimulating factor tyrosine kinase receptor; involved in regulation of growth and differentiation of myeloid cells	6.79
A1752668	Hs.78669	NNMT, nicotinamide N-methyltransferase	Nicotinamide N-methyltransferase; catalyzes the N-methylation of nicotinamide and other pyridines, structurally-related drugs and xenobiotics	6.52
AF182294	Hs.241578	LOC51691: U8 snRNA-associated Sm-like protein LSm8	Member of the Sm family; core constituent of snRNP complexes	6.50
AA457211	Hs.8858	BAZ1A, bromodomain adjacent to zinc finger domain, 1A	May bind DNA and act as a chromatin-mediated transcriptional regulator; contains a bromodomain and a PHD-finger	6.48
W40262	Hs.146310	zc7902.s1 Pancreatic Islet Homo sapiens cDNA clone IMAGE:328539 3', mRNA sequence	Function unknown	6.47
AB033091	Hs.74313	KIAA1285 protein	Function unknown	6.45
AA282898	Hs.163900	ESTs, Highly similar to winged helix/forKhead transcription factor [Homo sapiens] [H.sapiens]	Function unknown	6.36
BE613269	Hs.21893	DKFZp761N0824: hypothetical protein	Function unknown	6.35
H25836	Hs.301527	ESTs, Moderately similar to unknown [Homo sapiens] [H.sapiens]	Function unknown	6.27
AL037228	Hs.82043	NUDT5, nudix (nucleoside diphosphate linked moiety X)-type motif 5	NDP-sugar hydrolase; converts ADP-ribose to AMP or ribose 5-phosphate; contains a NudT motif	6.25
AV662037	Hs.124740	FLJ30532: hypothetical protein FLJ30532	Function unknown	6.21

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
AI874383	Hs.22891	wc38h08.x1 NCL CGAP_P728 Homo sapiens cDNA clone IMAGE:2320959 3', mRNA sequence	Function unknown	6.20
AW342140	Hs.182845	ESTs, Weakly similar to POL2_MOUSE Retrovirus-related POL polyprotein [Contains: Reverse transcriptase ; Endonuclease] [M.musculus]	Function unknown	6.18
BE580135	Hs.5232	HSPC125, HSPC125 protein	Function unknown	6.17
BE409857	Hs.69499	HSPC132: hypothetical protein HSPC132	Moderately similar to a region of <i>S. cerevisiae</i> Yki053c-ap	6.16
AW972542	Hs.288008	LLOC116150: hypothetical protein, MGC:7199	Function unknown	6.16
AI523755	Hs.69236	DKFZP434L0716: hypothetical protein DKFZP434L0718	Function unknown	6.16
NM_014058	Hs.7917	DKFZP564K247: DKFZP564K247 protein	Function unknown	6.08
AI857607	Hs.181301	CTSS, cathepsin S	Cathepsin S; lysosomal cysteine (thiol) protease that cleaves elastin	6.04
AW247529	Hs.8793	PAFAH1B3, platelet-activating factor acetylhydrolase, isoform lb, gamma subunit (28kD)	Platelet-activating factor acetylhydrolase gamma; may play a role in brain development	5.98
AK000868	Hs.5570	Homo sapiens cDNA FLJ10006 fis. clone HEMBA1000168, weakly similar to CYLCIN1	Function unknown	5.92
AF053551	Hs.31684	MTX2, metaxin 2	Very strongly similar to murine metaxin 2 (Mm.12941); are involved in mitochondrial protein import	5.91
AI538613	Hs.298241	TMPRSS3, Transmembrane protease, serine 3	The encoded protein contains a serine protease domain, a transmembrane domain, a LDL receptor-like domain, and a scavenger receptor cysteine-rich domain. This gene was identified as a tumor associated gene that is overexpressed in ovarian tumors.	5.86
U48508	Hs.89631	Human skeletal muscle ryanodine receptor gene (RYR1), exons 103, 104, 105, 106, and complete cds.	Function unknown	5.86
T69387	Hs.76364	AIF1, allograft inflammatory factor 1	Allograft inflammatory factor 1; cytokine inducible protein associated with vascular injury	5.86
AC005954	Hs.25527	Homo sapiens chromosome 19, cosmid R28784, complete sequence	Function unknown	5.86
AB037805	Hs.88442	KIAA1384 protein	Function unknown	5.84

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
AL031427	Hs.40094	Human DNA sequence from clone 167A19 on chromosome 1p32.1-33. Contains three genes for novel proteins, the DIO1 gene for type I iodothyronine deiodinase (EC 3.8.1.4, TXD1, ITD1) and an HNRNP A3 (Heterogenous Nuclear Ribonucleoprotein A3, FBRNP) pseudogene.	Function unknown	5.83
AA340884	Hs.278562	CLDN7, claudin 7	Similar to murine Cldn7; are an integral membrane protein	5.76
X89984	Hs.211563	BCL7A, B-cell CLL/lymphoma 7A	Similar to the actin-binding protein caldesmon; serine-rich	5.74
AI355761	Hs.242463	q19a11.x1 NCL CGAP_Co14 Homo sapiens cDNA clone IMAGE:1662908.3 similar to gb:X74929 KERATIN, TYPE II CYTOSKELETAL 8 (HUMAN); mRNA sequence	Function unknown	5.73
AA376409	Hs.10862	Homo sapiens cDNA: FLJ23313 fis, clone HEP11919	Function unknown	5.71
AA310182	Hs.169248	HCS: cytochrome c	Somatic cytochrome c (heart cytochrome c)	5.67
AW015534	Hs.217483	ANXA2, annexin A2	Annexin II (lipocortin-2); enhances osteoclast formation and bone resorption; member of the annexin protein family	5.64
AA326108	Hs.53631:82	BHLHB3; basic helix-loop-helix domain containing, class B, 3	Basic helix-loop-helix (bHLH) transcription factors (e.g., DEC1, also called BHLHB2; 604256) are related to Drosophila hairy/enhancer of split proteins. They are involved in the control of proliferation and development during differentiation, particularly in neurons.	5.64
AA120865	Hs.23136	ESTs, Highly similar to THYA_HUMAN Prothymosin alpha [H.sapiens]	Function unknown	5.62
AK000517	Hs.6844	NALP2; NALP2 protein	Protein with low similarity to murine Op1	5.54
Z36842	Hs.57548	H.sapiens (xs85) mRNA, 209bp	Function unknown	5.53
AA831552	Hs.268016	Homo sapiens cDNA: FLJ21243 fis, clone COL01164	Function unknown	5.50
AL137578	Hs.27607	Homo sapiens mRNA; cDNA DKFZp564N2464 (from clone DKFZp564N2464)	Function unknown	5.50
AA316181	Hs.61635	STEAP, six transmembrane epithelial antigen of the prostate	Six transmembrane epithelial antigen of the prostate; prostate-specific cell-surface antigen	5.46

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
X03635	Hs.1657	ESR1, estrogen receptor 1	Estrogen receptor; nuclear receptor transcription factor activated by ligand-binding. Involved in hormone-mediated inhibition of gene expression	5.42
AI557280	Hs.184270	PT2.1_15_G11.r tumor2 Homo sapiens cDNA 3', mRNA sequence	Function unknown	5.41
AW248508	Hs.278727	Homo sapiens cDNA FLJ14035 fis, clone HEMBA1004638	Function unknown	5.40
N90866	Hs.278770	CDW52, CDW52 antigen (CAMPATH-1 antigen)	CAMPATH-1 antigen; GPI-anchored protein	5.39
U83115	Hs.161002	AIM1, absent in melanoma 1	Member of the beta gamma-crystallin superfamily of proteins; interactions with the cytoskeleton	5.35
AB007860	Hs.12802	ODEF2, development and differentiation enhancing factor 2	GTPase-activating protein; interacts with members of the Arf and Src family	5.35
Z46223	Hs.176663	H.sapiens DNA for immunoglobulin G Fc receptor IIIB	Immunoglobulin G Fc receptor	5.31
BE284974	Hs.8586	TRIP13, thyroid hormone receptor interactor 13	Interacts with ligand binding domain of thyroid hormone receptor and with human papillomavirus type 18 (HPV16) E1	5.30
AA194422	Hs.22584	MYO6, myosin VI	Motor, hearing, myosin ATPase, structural protein. Class 8 myosin; motor protein; very strongly similar to murine Myo6	5.27
AF134157	Hs.169487	MAFB, v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian)	Very strongly similar to murine Krm1; may function as a basic domain-leucine zipper transcription factor	5.25
AA232119	Hs.16085	SH120; putative G-protein coupled receptor	putative G-protein coupled receptor	5.25
W58353	Hs.285123	OSBPL10, oxysterol binding protein-like 10	Member of the oxysterol-binding protein (OSBP) family; may bind oxygenated derivatives of cholesterol	5.21
AW167128	Hs.231934	ESTs, Weakly similar to A57717 transcription factor EC2 - human [H.sapiens]	Function unknown	5.19
U70370	Hs.84136	PITX1, paired-like homeodomain transcription factor 1	Member of the homeodomain family of DNA binding proteins; may regulate gene expression and control cell differentiation	5.18
N55669	Hs.333823	MRPL13, mitochondrial ribosomal protein L13	Protein of the large 60S ribosomal subunit	5.17
BE298446	Hs.305880	BCL2L1, BCL2-like 1	BCL2-related protein; alternative form bcl-xlong inhibits apoptosis and bcl-xshort induces apoptosis	5.17
AW136551	Hs.181245	Homo sapiens cDNA FLJ12532 fis, clone NT2RM4000200	Function unknown	5.15
AW250380	Hs.109059	HGS, hepatocyte growth factor-regulated tyrosine kinase substrate	Zinc-finger protein; interacts with STAM, undergoes tyrosine phosphorylation in response to IL2, CSF2, or HGF	5.13

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
AW002565	Hs.124860	Homo sapiens cDNA: FLJ21763 fis. clone COLF6967	Function unknown	5.13
A1697274	Hs.105435	GMD5, GDP-mannose 4,8-dehydratase	GDP-mannose-4,8-dehydratase; epimerase converts GDP-mannose to GDP-mannose-4-keto-6-D-deoxymannose, plays a role in the synthesis of fucosylated oligosaccharides	5.11
NM_003878	Hs.78819	GGH, gamma-glutamyl hydrolase (conjugase, folypolygammaglutamyl hydrolase)	Gamma-glutamyl hydrolase; has greater exopeptidase activity on methoxate pentaglutamates than on diglutamate	5.11
AF052112	Hs.12540	LYPLA1, lysophospholipase I	Lysophospholipid-specific lysophospholipase 1; hydrolyzes lysophosphatidyl choline	5.09
AV654694	Hs.82316	IFI44, interferon-induced protein 44	Member of the family of Interferon-alpha/beta inducible proteins; may mediate the antiviral action of Interferon	5.09
R24801		Homo sapiens adenylosuccinate synthetase isozyme (ADSS) mRNA, complete cds	Adenylosuccinate synthetase	5.07
BE016020	Hs.85838	Homo sapiens cDNA clone IMAGE:2983945 6' similar to TR:O15427 O15427 MONOCARBOXYLATE TRANSPORTER, 11, mRNA sequence	Function unknown	5.04
AW163799	Hs.198365	BPGM, 2,3-bisphosphoglycerate mutase	2,3-bisphosphoglycerate mutase; has synthase, mutase, and phosphatase activities, controls 2,3-diphosphoglycerate metabolism, which is an effector for haemoglobin	5.04
AA278821	Hs.1908	PRG1, proteoglycan 1, secretory granule	Secretory granule proteoglycan 1	5.02
NM_003726	Hs.10126	SCAP1, src family associated phosphoprotein 1	Src kinase-associated phosphoprotein; acts as an adaptor protein; contains a pleckstrin homology domain and an SH3 domain	5.02
AA281167	Hs.111911	ESTs, Weakly similar to T08291 extensin homolog T8E8.80 - Arabidopsis thaliana [A.thaliana]	Function unknown	5.02
		C90003067:gi12737280 ref XP_006682.2  keratin 18 [Homo sapiens] j8633	Function unknown	5.01
AF098158	Hs.9329	C20orf1, chromosome 20 open reading frame 1	Proliferation-associated nuclear protein; associates with the spindle pole and mitotic spindle during mitosis	5.00

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
AA101043	Hs.151254:19	KLK7, kallikrein 7 (chymotryptic; stratum corneum)	Epidermal differentiation. Stratum corneum chymotryptic enzyme; serine protease. Growing evidence suggests that many kallikreins are implicated in carcinogenesis and some have potential as novel cancer and other disease biomarkers. Thought to be involved in the proteolysis of intercellular cohesive structures preceding desquamation, which is the shedding of the outermost layer of the epidermis. Function unknown	4.87
AF017986	Hs.31366:185	Homo sapiens secreted apoptosis related protein 1 (SARP-1) mRNA, partial cds.		4.12
AW680584	Hs.3337:137	TM4SF1: transmembrane 4 superfamily member 1	Pathogenesis, plasma membrane, cell proliferation, N-linked glycosylation, integral membrane protein, integral plasma membrane protein, L6 antigen: member of the transmembrane 4 superfamily (TM4SF). The proteins mediate signal transduction events that play a role in the regulation of cell development, activation, growth and motility. This encoded protein is a cell surface antigen and is highly expressed in different carcinomas.	3.62
W28092	Hs.7678:40	CRABP1 Cellular retinoic acid binding protein 1	Cytoplasm, retinoid binding, signal transduction, developmental processes. Cellular retinoic acid-binding protein 1: are involved in delivering retinoic acid to the nucleus, assumed to play an important role in retinoic acid-mediated differentiation and proliferation processes.	3.34
HP3366	Hs.7587:84	Homo sapiens cDNA: FLJ121962 fis, clone HEP05584	Function unknown	3.29
D49441	Hs.155981:53	MSLN, mesothelin	Cell adhesion, cell surface antigen, membrane. Pre-pro-megakaryocyte potentiating factor. An antibody that reacts with ovarian cancers and mesotheliomas was used to isolate a cell surface antigen named mesothelin. Although the function of mesothelin is unknown, it may play a role in cellular adhesion and is present on mesothelium, mesotheliomas, and ovarian cancers.	3.14
AA214228	Hs.127751:21,Hs.78006:5	C20orf180: chromosome 20 open reading frame 180	Region of high similarity to tyrosine-phosphorylated protein DOK1	2.99
M31126	Hs.272620:1	PSG9: pregnancy specific beta-1-glycoprotein 9	Pregnancy, extracellular, plasma glycoprotein. Member of the pregnancy-specific glycoprotein (PSG) and CEA families.	2.82
U82801	Hs.79381:85	KLK8, kallikrein 8 (neurosin, zyme)	Serine type peptidase, pathogenesis. Neurosin (protease M, zyme); a serine protease that cleaves amyloid precursor protein (APP). Growing evidence suggests that many kallikreins are implicated in carcinogenesis and some have potential as novel cancer and other disease biomarkers.	2.77
AK001638	Hs.285803:8	Homo sapiens cDNA FLJ12852 fis, clone NT2RP2003445	Function unknown	2.73

Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
NM_014767	Hs.74583:140	KIAA0275: KIAA0275 gene product	Function unknown	2.72
NM_000699	Hs.75733:128,Hs.278398:100,Hs.274376:1	AMY2A: amylase, alpha 2A; pancreatic	Alpha-amylase, extracellular space, carbohydrate metabolism. Pancreatic alpha-amylase 2A (1,4-alpha-D-glucan glucanohydrolase); cleaves internal $\alpha$ -1,4 bonds between glucose monomers to digest starch.	2.71
AA430348	Hs.288837:40	Homo sapiens cDNA FLJ12927 fis, clone NT2RP2004743	Function unknown	2.69
X51830	Hs.1145:22,Hs.296851:1	WT1, Wilms tumor 1	Nucleus, transcription factor, transcription regulation. 4 Zn finger domains. Functions in kidney and gonad proliferation and differentiation. Mutations in this gene are associated with the development of Wilms tumors in the kidney or with abnormalities of the genitourinary tract.	2.58
BE393948	Hs.50915:17	KLK6, kallikrein 5	Serine type peptidase, epidermal differentiation, extracellular space. Stratum corneum tryptic enzyme (kallikrein-like protein); may function in epidermal stratum corneum desquamation and turnover. Expression in prostate cancer negatively correlated with cancer aggressiveness (Yousef 2002)	2.34
NM_002776	Hs.88423:48	KLK10, kallikrein 10	Putative serine protease. Expressed in normal breast tissue and benign lesions, with loss of expression during tumor progression (Dhar 2001). SNPs associated with prostate, breast, testicular, and ovarian cancers (Bhargava 2002).	2.24
NM_000854	Hs.8272:284	PTGS2: prostaglandin D2 synthase (21kD, brain)	Membrane, prostaglandin-D synthase. Glutathione-independent prostaglandin D2 synthase; membrane associated, catalyzes synthesis of prostaglandin D; member of the lipocalin family of transporters.	2.15
AB029000	Hs.70823:109,Hs.297970:48	KIAA1077: sulfatase FP	Function unknown	2.04
AL044315	Hs.173094:70	KIAA1750: KIAA1750 protein	Function unknown	0.95
AA334592	Hs.79814:337	LUM: lumican	Vision, proteoglycan, extracellular matrix, cartilage condensation, extracellular matrix glycoprotein. Member of the specialized collagens and SLRP family	0.93
S78895	Hs.83942:248	CTSK: cathepsin K (pseudodysostosis)	Lysosome, cathepsin K, cysteine-type peptidase, proteolysis and peptidolysis. Cathepsin K (cathepsin O), a cysteine (thiol) protease; involved in bone remodeling and reabsorption	0.91
AI091185	Hs.65029:120	Homo sapiens cDNA clone IMAGE:1568910 3', mRNA sequence	Function unknown	0.91



Accession number	UniGene Mapping	Gene symbol and title	Putative Function	Ratio
AF026692; NM_003014	Hs.105700:83;Hs.278611:3	SFRP4: secreted frizzled-related protein 4	Member of the SFRP family that contains a cysteine-rich domain homologous to the putative Wnt-binding site of Frizzled proteins. SFRPs act as soluble modulators of Wnt signaling. The expression of SFRP4 in ventricular myocardium correlates with apoptosis related gene expression. Function unknown	0.73
A1893243	Hs.97258:15	ESTs, Moderately similar to S29539 ribosomal protein L13a, cytosolic	Function unknown	-2.98
A1287700	Hs.111128:7	Homo sapiens, clone IMAGE:4108329, mRNA	Function unknown	-5.71
AA291377	Hs.50831:23	Homo sapiens Ly-6 antigen/CD4 receptor-like domain-containing protein mRNA, complete cds	Function unknown	-6.78
A1420213	Hs.149722:3	cDNA clone IMAGE:2094208 3', mRNA sequence	Function unknown	-8.52
AJ245871	Hs.12844:73	EGFL6, EGF-like-domain; multiple 6	Cell cycle, oncogenesis, integrin ligand, extracellular space. Member of the epidermal growth factor (EGF) repeat superfamily; contains an EGF-like-domain. Expressed early during development, and its expression has been detected in lung and meningioma tumors.	-9.44
AB018305	Hs.5378:149	SPON1, spondin 1, (F-spondin) extracellular matrix protein	Extracellular matrix protein. Very strongly similar to rat F-spondin (Rn.7546); may have a role in the growth and guidance of axons.	-12.55
AW872527	Hs.59761:19	ESTs; Weakly similar to DAP1_HUMAN DEATH-ASSOCIATED PROTEIN 1	Function unknown	-14.17
AF128755	Hs.117772:9;Hs.88474:1	Homo sapiens prostaglandin endoperoxide H synthase-1 mRNA, partial 3' untranslated region.	Function unknown	-21.34
A1023799	Hs.163242:5	Homo sapiens cDNA clone IMAGE:1655725 3' similar to contains MER20.12 MER20 repetitive element 1, mRNA sequence	Function unknown	-41.34

Table 3  
Preferred diagnostic and prognostic markers for detecting ovarian cancer or a recurrence thereof  
or survival of a subject suffering from ovarian cancer

A. DOWN-REGULATED GENES			Function	SEQ ID NO:	Chromosome Location	P value
Accession Number	Unigene Mapping	Gene Name				
AI631024; NM_005460	Hs.24848:32; Hs.300445:4	SNCAIP, alpha synuclein, alpha interacting protein (synphilin)	Cytoplasm, pathogenesis, protein binding. Synphilin-1; promotes formation of cytosolic inclusions in neurons (SNCAIP). Synuclein alpha interacting protein contains several protein-protein interaction domains and interacts with alpha synuclein in neurons. Mutations of SNCAIP have been linked to Parkinson disease.	SEQ ID NO: 1 (DNA) SEQ ID NO: 2 (PRT)	5q23.2	0
NM_002387	Hs.1345:5	MCC, mutated in colorectal cancers	Receptor, signal transduction, tumor suppressor. Similar to the G protein-coupled m3 muscarinic acetylcholine receptor. MCC is a candidate for the putative colorectal tumor suppressor gene. The MCC gene product are involved in early stages of colorectal neoplasia in both sporadic and familial tumors.	SEQ ID NO: 3 (DNA) SEQ ID NO: 4 (PRT)	5q22.2	0
AI420582; NM_022117	Hs.136164:23	SE20-4, cutaneous T-cell lymphoma- associated tumor antigen se20- 4se20-4	Cutaneous T-cell lymphoma-associated tumor antigen se20-4se20-4; differentially expressed nucleolar TGF- beta1 target protein (DENTT); also known as CDA1	SEQ ID NO: 5 (DNA) SEQ ID NO: 6 (PRT)	unmapped	0

Table 3 continued

Accession Number	Unigene Mapping	Gene Name	Function	SEQ ID NO:	Chromosome Location	P value
<b>B. UP-REGULATED GENES</b>						
BC006428; NM_016463	Hs.15093:210;Hs.290304:1	HSPC195, hypothetical protein HSPC195 FLJ20171, hypothetical protein FLJ20171 MAL2	Homo sapiens cDNA FLJ10920 fis, clone OVARC1000384-resourcerer.	SEQ ID NO: 7 (DNA) SEQ ID NO: 8 (PRT)	5q31.2	0
NM_017697	Hs.24743:84		contains 3 RNA recognition motifs	SEQ ID NO: 9 (DNA) SEQ ID NO: 10 (PRT)	8q22.1	0
AW630088; NM_001306	Hs.76550:164		Mal2 T-cell differentiation protein; found thru interaction with TP52 which is overexpressed in breast cancer, 4 TM are involved in vesicle transport	SEQ ID NO: 11 (DNA) SEQ ID NO: 12 (PRT)	8q24.12	0
NM_015238	Hs.21543:36	KJAA0869, KIAA0869 protein; KIBRA	Function unknown	SEQ ID NO: 13 (DNA) SEQ ID NO: 14 (PRT)	5q34	0.0002
AA284879	Hs.25840:284;Hs.5372:2	CLDN3, claudin 3	Integral plasma membrane protein, pathogenesis, tight junction, transmembrane receptor. Member of the claudin family of integral membrane proteins; receptor for Clostridium perfringens enterotoxin;	SEQ ID NO: 15 (DNA) SEQ ID NO: 16 (PRT)	7q11.23	0.0004
NM_022454	Hs.97984:22	SOX17, SRY (sex determining region Y)-box 17	Likely ortholog of mouse SRY-box containing gene 17; alias SOX17	SEQ ID NO: 17 (DNA) SEQ ID NO: 18 (PRT)	8q11.23	0.0005
NM_005682	Hs.6527:201	GPR56, G protein-coupled receptor 56	cell adhesion, cell-cell signalling, G-protein linked receptor, integral plasma membrane protein, G-protein linked receptor protein signalling pathway. Member of the G protein-coupled receptor family; similar to secretin and calcitonin receptors. 7 transmembrane domains, a much-like domain and cysteine box in the N-terminal region. Expressed in range of tissues, highest levels in thyroid, selectively within the monolayer of cuboidal epithelial cells of the smaller, more actively secreting follicles of human thyroid. Differentially expressed in melanoma cell lines with different metastatic potential (Zendman et al 1999).	SEQ ID NO: 19 (DNA) SEQ ID NO: 20 (PRT)	18q13	0.0012
NM_001307	Hs.276562:101	CLDN7, claudin 7	Integral membrane protein, tight junction. Similar to murine Cldn7;	SEQ ID NO: 21 (DNA) SEQ ID NO: 22 (PRT)	17p13.1	0.0016

NM_014736	Hs.81892:95	KIAA0101 gene product	function unknown; no significant hits with Superfamily	SEQ ID NO: 23 (DNA) SEQ ID NO: 24 (PRT)	15q31	0.0025
BE184455; NM_003064	Hs.251754:128.H s.245742:1	SLPI, secretory leukocyte protease inhibitor (antileukoproteina se)	Plasma protein, proteinase inhibitor. Secreted inhibitor which protects epithelial tissues from serine proteases. Found in various secretions including seminal plasma, cervical mucus, and bronchial secretions, has affinity for trypsin, leukocyte elastase, and cathepsin G. Its inhibitory effect contributes to the immune response by protecting epithelial surfaces from attack by endogenous proteolytic enzymes; the protein is also thought to have broad-spectrum anti-biotic activity.	SEQ ID NO: 25 (DNA) SEQ ID NO: 26 (PRT)	20q13.12	0.0034
NM_013994	Hs.75562:147	DDR1, discoidin domain receptor family, member 1	Cell adhesion, integral plasma membrane protein, transmembrane receptor, protein tyrosine kinase. Epithelial-specific receptor protein tyrosine kinase; are involved in cell adhesion; has putative discoidin motifs in extracellular domain. DDR1 (CD167a) is a RTK that is widely expressed in normal and transformed epithelial cells and is activated by various types of collagen.	SEQ ID NO: 27 (DNA) SEQ ID NO: 28 (PRT)	6p21.33	0.0055
NM_001067	Hs.156348:184.H s.270810:2	TOP2A, topoisomerase (DNA) II alpha (170kD)	DNA binding, DNA topoisomerase (ATP-hydrolyzing), nucleus. DNA topoisomerase II alpha; may relax DNA torsion upon replication or transcription. Involved in processes such as chromosome condensation, chromatid separation, and the relief of torsional stress that occurs during DNA transcription and replication. Catalyzes the transient breaking and rejoining of two strands of duplex DNA. The gene encoding this enzyme functions as the target for several anticancer agents and a variety of mutations in this gene have been associated with the development of drug resistance. Reduced activity of this enzyme may also play a role in ataxia-telangiectasia.	SEQ ID NO: 29 (DNA) SEQ ID NO: 30 (PRT)	17q21.2	0.006

BE386983; NM_138410	Hs.343214	CKLFSF7; chemokine-like factor super family 7	chemokine-like factor gene superfamily; transmb 4 superfamily	SEQ ID NO: 31 (DNA) SEQ ID NO: 32 (PRT)	3p23	0.0131
AF098158; NM_012112	Hs.9329:152	C20orf1, chromosome 20 open reading frame 1	ATP binding, GTP binding, cell proliferation, mitosis, nucleus spindle. Proliferation-associated nuclear protein; associates with the spindle pole and mitotic spindle during mitosis	SEQ ID NO: 33 (DNA) SEQ ID NO: 34 (PRT)		0.0183
NM_001769	Hs.1244:227,Hs.2 30559:1,Hs.2420 20:1	CD9; CD9 antigen (p24)	Plasma membrane, integral plasma membrane protein, Member of the transmembrane 4 superfamily (TM4SF); may mediate platelet activation and aggregation. Cell surface glycoprotein that is known to complex with integrins and other transmembrane 4 superfamily proteins.	SEQ ID NO: 35 (DNA) SEQ ID NO: 36 (PRT)	12p13.31	0.0008
NM_020859	Hs.278628:52	ShrmL, Shroom- related protein (KIAA1481 protein)	Amliloride-sensitive sodium channel (weakly similar to Mus musculus PDZ domain actin binding protein)	SEQ ID NO: 37 (DNA) SEQ ID NO: 38 (PRT)		0.0074
NM_004433	Hs.166096:170	ELF3, E74-like factor 3 (ets domain transcription factor, epithelial- specific)	Embryogenesis and morphogenesis, transcription co- activator, transcription factor, transcription from Pol II promoter. ETS domain transcriptional activator; activates expression of epithelial cell specific genes.	SEQ ID NO: 39 (DNA) SEQ ID NO: 40 (PRT)	1q32.1	0.0004
AI791905; NM_019027	Hs.95549:147,Hs. 229556:1	FLJ20273, RNA- binding protein	Contains four RNA recognition motifs (RRM, RBD, or RNP)	SEQ ID NO: 41 (DNA) SEQ ID NO: 42 (PRT)		0.0007
X69699; NM_013952	Hs.73149:72,Hs.2 13008:1	PAX8, paired box gene 8	Histogenesis and organogenesis, embryogenesis and morphogenesis, thyroid-stimulating hormone receptor, transcription factor. Member of the paired domain family of nuclear transcription factors; are involved in the ribosome assembly, required for normal thyroid development. PAX genes play critical roles during fetal development and cancer growth.	SEQ ID NO: 43 (DNA) SEQ ID NO: 44 (PRT)	2q13	0.0009
AI301558	Hs.280801:35, Hs.356228	EST	Function unknown	SEQ ID NO: 45 (DNA)		0.0044
NM_018000	Hs.79741:18	FLJ10116, hypothetical protein FLJ10116	Function unknown	SEQ ID NO: 46 (DNA) SEQ ID NO: 47 (PRT)	2q35	0.0051
NM_144724	Hs.124740:18	hypothetical protein FLJ30532	59% identity to human Zinc finger protein 91	SEQ ID NO: 48 (DNA) SEQ ID NO: 49 (PRT)	5q13.12	0.0051

AF111856; NM_006424	Hs.105039:48	SLC34A2, solute carrier family 34 (sodium phosphate), member 2	SLC34A2: solute carrier family 34 (sodium phosphate), member 2; contains 8 predicted TMs and a cysteine-rich N-terminal region. Type 2 sodium-dependent phosphate transporter, member of the renal type II co-transporter family.	SEQ ID NO: 50 (DNA) SEQ ID NO: 51 (PRT)	4p15.2	0.0121
AW959311	Hs.87019:8; Hs.172012	EST DKFZp434J037	probable serine/threonine protein kinase; KIAA0537	SEQ ID NO: 52 (DNA)	1q32.1	0.0251
AF111713	Hs.286218:64	JAM1, junctional adhesion molecule	Cell motility, inflammatory response, intercellular junction. Role in the regulation of tight junction assembly in epithelia. Ligand of JAM is required for reovirus-induced activation of NF-kappa-B and apoptosis. Role in lymphocyte homing.	SEQ ID NO: 53 (DNA) SEQ ID NO: 54 (PRT)		0.0261
AJ078811; NM_006638	Hs.154872:123	MTHFD2, methylene tetrahydrofolate dehydrogenase (NAD+ dependent); methylenetetrahydrofolate	Electron transporter, methylenetetrahydrofolate cyclohydrolase, mitochondrial. encodes a nuclear-encoded mitochondrial bifunctional enzyme with methylenetetrahydrofolate dehydrogenase and methylenetetrahydrofolate cyclohydrolase activities. may provide formyltetrahydrofolate for formylmethionyl tRNA synthesis; involved in initiation of mitochondrial protein synthesis.	SEQ ID NO: 55 (DNA) SEQ ID NO: 56 (PRT)	2p13.1	0.0342

Table 3 continued

Accession Number	Unigene Mapping	Gene Name	Function	SEQ ID NO:	Chromosome Location	P value
AA584890; NM_006149	Hs.5302:132	LGALS4, lectin, galactoside-binding, soluble, 4 (galectin 4)	Lectin, cytosol, cell adhesion, plasma membrane. Binds to beta galactoside, involved in cell adhesion, cell growth regulation, inflammation, immunomodulation, apoptosis and metastasis; member of a family of lectins. LGALS4 is an S-type lectin that is strongly underexpressed in colorectal cancer.	SEQ ID NO: 57 (DNA) SEQ ID NO: 58 (PRT)		0.0001
	Hs.89436:50	CDH17, cadherin 17, LI cadherin (liver-intestine)	Cell adhesion, integral plasma membrane protein, membrane fraction, small molecule transport, transporter. Member of the cadherin family of calcium-dependent glycoproteins; facilitates uptake of peptide-based drugs, may mediate cell-cell interactions. Component of the gastrointestinal tract and pancreatic ducts, intestinal proton-dependent peptide transporter in the first step in oral absorption of many medically important peptide-based drugs.	SEQ ID NO: 59 (DNA) SEQ ID NO: 60 (PRT)	19q13.2	0.0172
NM_005588	Hs.179704	MEP1A, meprin A alpha, PABA peptide hydrolase	metalloprotease located apically and secreted by epithelial cells in normal colon; degrades broad range of ECM components in vitro; proposed role in tumour progression by facilitating migration, intravasation and metastasis	SEQ ID NO: 61 (DNA) SEQ ID NO: 62 (PRT)	8q22.1	0.01
					6p12	

Table 3 continued

Accession Number	Unigene Mapping	Gene Name	Function	SEQ ID NO:	Chromosome Location	P value
<b>D. PROGNOSTIC MARKERS FOR SURVIVAL OR RECURRENCE</b>						
NM_015092	Hs.278428	DD5; EDD	Homo sapiens progesterin induced protein (DD5), mRNA. EDD; Soluble fraction, cell proliferation, ubiquitin-protein ligase, ubiquitin conjugating enzyme, ubiquitin-dependent protein degradation. Member of the HECT family of proteins; may function as an E3 ubiquitin-protein ligase. This gene is localized to chromosome 8q22, a locus disrupted in a variety of cancers. This gene potentially has a role in regulation of cell proliferation or differentiation.	SEQ ID NO: 63 (DNA) SEQ ID NO: 64 (PRT)		0.00
BE465867; NM_014992	Hs.197751:66	DAAM1	dishevelled associated activator of morphogenesis 1 The protein encoded by this gene contains FH domains and belongs to a novel FH protein subfamily implicated in cell polarity, thought to function as a scaffolding protein.	SEQ ID NO: 65 (DNA) SEQ ID NO: 66 (PRT)	8q22.3	0.04
AA381553; NM_002122	Hs.198253:21	HLA1QA	major histocompatibility complex, class II, DQ alpha 1 Pathogenesis, class II major histocompatibility complex antigen. Alpha 1 chain of HLA-DQ1 class II molecule (a antigen); complex binds peptides and presents them to CD4+ T lymphocytes Proteome	SEQ ID NO: 67 (DNA) SEQ ID NO: 68 (PRT)	14q23.1	0.00
AF026692; NM_003014	Hs.105700:83,Hs.278611:3	SFRP4; secreted frizzled-related protein 4	Member of the SFRP family that contains a cysteine-rich domain homologous to the putative Wnt-binding site of Frizzled proteins. SFRPs act as soluble modulators of Wnt signaling. The expression of SFRP4 in ventricular myocardium correlates with apoptosis related gene expression.	SEQ ID NO: 69 (DNA) SEQ ID NO: 70 (PRT)	6p21.3 7p14	0.73
AW015534; NM_004039	Hs.217493	ANXA2, annexin A2	Annexin II (lipocortin-2); enhances osteoclast formation and bone resorption; member of the annexin protein family	SEQ ID NO: 71 (DNA) SEQ ID NO: 72 (PRT)	15q21-22	0.00
BE24669; NM_003955	Hs.345728	SOCS3	STAT induced STAT-inhibitor 3; suppressor of cytokine signalling 3; suppression of IL-6 mediated signalling	SEQ ID NO: 73 (DNA) SEQ ID NO: 74 (PRT)	17q25.3	0.02
AI677897; NM_014059	Hs.76640	RGC32	RGC32, hypothetical protein, unknown function	SEQ ID NO: 75 (DNA) SEQ ID NO: 76 (PRT)	13q13.3	0.04



AA829286; NM_000331	Hs.332053	SAA1, serum amyloid A1 FLJ10134, hypothetical protein	Serum amyloid A1; high density lipoprotein; role in cholesterol metabolism; inflammatory response	SEQ ID NO: 77 (DNA) SEQ ID NO: 78 (PRT)	11p15.1	0.04
AA243499; NM_018004	Hs.104800	GJB2, gap junction protein beta2; connexin 28	Unknown	SEQ ID NO: 79 (DNA) SEQ ID NO: 80 (PRT)	3q12.3	0.01
M86849; NM_004004	Hs.323733	NOV1; Nephroblastoma overexpressed gene	Cellular gap junctions; mutations cause some forms of deafness	SEQ ID NO: 81 (DNA) SEQ ID NO: 82 (PRT)	13q11-12	0.00
NM_002514	Hs.235935		Role in cell adhesion and migration in endothelial cells; promotes cell survival	SEQ ID NO: 83 (DNA) SEQ ID NO: 84 (PRT)	8q24.1	0.01

Table 4

Correlation of expression between normal ovarian surface epithelium (OSE), non-invasive tumors (borderline, BL) and ovarian cancer (CA) as determined by ANOVA

	CA125	MUC-1	E-cadherin	CLDN3	Ep-CAM	SOX17
OSE vs IC	<0.0001	<0.0001	0.7251	0.6132	0.1573	0.0854
OSE vs. BL	0.1765	<0.0001	0.0307	0.3633	0.0005	0.2287
OSE vs. CA	0.5443	<0.0001	0.1687	0.0008	<0.0001	0.6900
IC vs. BL	<0.0001	<0.0001	0.1116	0.7849	0.0913	0.2530
IC vs. CA	<0.0001	0.2707	0.4147	0.0071	0.0002	0.0544
BL vs. CA	0.0001	<0.0001	0.0615	<0.0001	0.0011	0.0152

5

Table 5

Correlation of gene expression with patient outcome (univariate analysis ie., expression alone without the influence of covariates)

Univariate analysis for clinicopathological variables and CLDN3, Ep-CAM, SOX17, CA125, MUC1 and E-cadherin immunoreactivity with survival and relapse in 156 patients with epithelial ovarian cancer

Variable	Disease Specific Survival		Relapse Free Survival	
	Univariate Hazards ratio (95% CI)	p-value	Univariate Hazards ratio (95% CI)	p-value
Pathological tumor stage				
Stage 1 - 3b vs. 3c - 4b	5.89 (3.214-10.79)	<0.0001	7.37 (3.26-16.63)	<0.0001
Tumor grade				
BL and G1 vs. G2 and G3	5.508 (2.745-11.052)	<0.0001	7.02 (2.76-17.82)	<0.0001
Age				
<50 vs. ≥50	0.533 (0.288-0.988)	0.0458	0.62 (0.29-1.33)	0.2221
Residual Disease				
RD<1cm vs. ≥1cm	4.192 (2.671-6.580)	<0.0001	4.17 (2.30-7.55)	<0.0001
CA125 level at diagnosis				
CA125 <500 vs. ≥500 U/ml	1.843 (1.102-3.080)	0.0197	2.292 (1.19-4.40)	0.0128
Performance Status				
PS<1 vs. ≥1	0.270 (0.133-0.549)	0.0003	0.53 (0.16-1.74)	0.2965
CLDN3 expression				
Membranous Score 0 vs. >0	2.794 (1.012-7.718)	0.0474	2.521 (0.908-6.998)	0.0758
Membranous Score <1 vs. ≥1	1.309 (0.763-2.246)	0.3285	1.952 (1.103-3.457)	0.0217
Ep-CAM expression				
Membranous Score <1 vs. ≥1	1.460 (0.809-2.634)	0.2093	2.041 (0.997-4.177)	0.0509
Membranous Score <2 vs. ≥2	1.041 (0.634-1.711)	0.873	1.449 (0.845-2.487)	0.1779
SOX17 expression				
Nuclear membranous Score 0 vs. >0	0.839 (0.514-1.368)	0.481	1.311 (0.728-2.358)	0.3667
Nuclear membranous Score <1 vs. ≥1	1.407 (0.615-3.218)	0.4183	1.037 (0.380-2.829)	0.9437
CA125 expression				
Membranous apical Score 0 vs. >0	2.581 (1.393-4.781)	0.0026	2.725 (1.218-6.093)	0.0146
Membranous apical Score <1 vs. ≥1	1.637 (1.045-2.564)	0.0313	1.298 (0.731-2.307)	0.3737
MUC1 expression				
Membranous apical Score 0 vs. >0	2.479 (0.343-17.898)	0.368	NA	
Membranous apical Score <1 vs. ≥1	3.745 (1.176-11.926)	0.0254	6.432 (1.562-26.483)	0.0099
Membranous apical Score <2 vs. ≥2	1.814 (0.898-3.664)	0.0969	3.893 (1.552-9.766)	0.0038
E-cadherin expression				
Membranous Score 0 vs. >0	0.806 (0.493-1.318)	0.3892	0.837 (0.477-1.467)	0.5341
Membranous Score <1 vs. ≥1	1.331 (0.532-3.333)	0.5411	0.847 (0.263-2.731)	0.7814
Membranous Score <2 vs. ≥2	0.593 (0.082-4.284)	0.6041	0.913 (0.125-6.646)	0.9284

Table 6  
Correlation of gene expression with patient outcome (multivariate  
analysis is looking at expression incorporating the influence of covariates)

5

Multivariate analysis for univariate significant clinicopathological variables and CLDN3, Ep-CAM, SOX17, CA125, MUC1  
and E-cadherin immunoreactivity with survival and relapse in 156 patients with epithelial ovarian cancer

Variable	Disease Specific Survival		Relapse Free Survival	
	Multivariate Hazards ratio (95% CI)	p-value	Univariate Hazards ratio (95% CI)	p-value
Pathological tumor stage				
Stage 1 - 3b vs. 3c - 4b	5.66 (2.467-13.012)	<0.0001	5.192 (1.860-14.496)	0.0017
Tumor grade				
BL and G1 vs. G2 and G3	4.919 (2.080-11.633)	0.0003	7.989 (2.385-26.760)	0.0008
Age				
<50 vs. ≥50	0.951 (0.482-1.877)	0.8853		
Residual Disease				
RD<1cm vs. ≥1cm	2.974 (1.783-4.959)	<0.0001	2.779 (1.433-5.393)	0.0025
CA125 level at diagnosis				
CA125 <500 vs. ≥500 U/ml	1.148 (0.625-2.109)	0.6563	1.289 (0.659-2.520)	0.4587
Performance Status				
PS<1 vs. ≥1	0.286 (0.136-0.601)	0.0009		
CLDN3 expression				
Membranous Score 0 vs. >0	1.165 (0.325-4.183)	0.8145		
Membranous Score <1 vs. ≥1			0.953 (0.473-1.919)	0.8918
CA125 expression				
Membranous apical Score 0 vs. >0	0.917 (0.415-2.025)	0.8302	0.693 (0.271-1.768)	0.4427
Membranous apical Score <1 vs. ≥1	1.664 (0.976-2.837)	0.0612		
MUC1 expression				
Membranous apical Score 0 vs. >0				
Membranous apical Score <1 vs. ≥1	0.678 (0.255-1.804)	0.4361		
Membranous apical Score <2 vs. ≥2				

## WE CLAIM:

1. A method of detecting an ovarian cancer-associated transcript in a biological sample, the method comprising contacting the biological sample with a polynucleotide  
5 that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Table 1 or 2 or 3.
2. A method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being  
10 tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein a modified level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said nucleic acid probe comprises a  
15 sequence selected from the group consisting of:
  - (i) a sequence comprising at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 46, 48, 50, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
  - 20 (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 46, 48, 50, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
  - 25 (iii) a sequence that is at least about 80% identical to a sequence selected from the group consisting of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 46, 48, 50, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
  - (iv) a sequence that encodes an amino acid sequence selected from the group  
30 consisting of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 47, 49, 51, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84; and
  - (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

3. A method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein a modified level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:

- 10 (i) a sequence comprising at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1, 5, 7, 9, 11, 13, 15, 17, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 45, 46, 48, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- 15 (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1, 5, 7, 9, 11, 13, 15, 17, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 45, 46, 48, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- 20 (iii) a sequence that is at least about 80% identical to a sequence selected from the group consisting of SEQ ID NOs: 1, 5, 7, 9, 11, 13, 15, 17, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 45, 46, 48, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- 25 (iv) a sequence that encodes an amino acid sequence selected from the group consisting of SEQ ID NOs: 2, 6, 8, 10, 12, 14, 16, 18, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 47, 49, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84; and
- (v) a sequence that is complementary to (i) or (ii) or (iii) or (iv).

4. The method of claim 2 or 3 wherein the hybridization is enhanced in the sample from the subject being tested compared to the hybridization obtained for a sample from a control subject not having ovarian cancer.

5. The method of claim 2 or 3 wherein the hybridization is reduced in the sample from the subject being tested compared to the hybridization obtained for a sample from a control subject not having ovarian cancer.

6. A method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein an enhanced level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian ovarian cancer, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:
- (i) a sequence comprising at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 or 2 other than a nucleic acid having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200;
  - (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 or 2 other than a nucleic acid having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200;
  - (iii) a sequence that is at least about 80% identical to (i) or (ii);
  - (iv) a sequence that encodes a polypeptide encoded by a nucleic acid set forth in Table 1 or 2 other than a nucleic acid having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200; and
  - (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).
7. The method of claim 6 wherein the nucleic acid probe comprises a sequence selected from the group consisting of:
- (i) a sequence comprising at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 7, 9, 11, 13, 15, 17, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 45, 46, 48, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
  - (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 7, 9, 11, 13, 15, 17, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 45, 46, 48, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;

- (iii) a sequence that is at least about 80% identical to a sequence selected from the group consisting of SEQ ID NOs: 7, 9, 11, 13, 15, 17, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 45, 46, 48, 52, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- 5 (iv) a sequence that encodes an amino acid sequence selected from the group consisting of SEQ ID NOs: 8, 10, 12, 14, 16, 18, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 47, 49, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84; and
- (v) a sequence that is complementary to any one of the sequences set forth in (i) or
- 10 (ii) or (iii) or (iv).

8. A method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for

15 hybridization to occur and then detecting the hybridization wherein a reduced level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian ovarian cancer, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:

- 20 (i) a sequence comprising at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200;
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1
- 25 and having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200;
- (iii) a sequence that is at least about 80% identical to (i) or (ii);
- (iv) a sequence that encodes a polypeptide encoded by a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of
- 30 NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200; and
- (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

9. The method of claim 8 wherein the nucleic acid probe comprises a sequence

35 selected from the group consisting of:

- (i) a sequence comprising at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1, 3, and 5;
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1, 3, and 5;
- (iii) a sequence that is at least about 80% identical to a sequence selected from the group consisting of SEQ ID NOs: 1, 3, and 5;
- (iv) a sequence that encodes an amino acid sequence selected from the group consisting of SEQ ID NOs: 2, 4, and 6; and
- (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).
10. The method according to any one of claims 1 to 9 wherein the ovarian cancer that is diagnosed is an epithelial ovarian cancer.
11. The method according to any one of claims 1 to 9 wherein the ovarian cancer that is diagnosed is selected from the group consisting of serous ovarian cancer, non-invasive ovarian cancer, mixed phenotype ovarian cancer, mucinous ovarian cancer, endometrioid ovarian cancer, clear cell ovarian cancer, papillary serous ovarian cancer, Brenner cell and undifferentiated adenocarcinoma.
12. The method according to claim 11 wherein the ovarian cancer that is diagnosed is selected from the group consisting of serous ovarian cancer, mucinous ovarian cancer, endometrioid ovarian cancer and clear cell ovarian cancer.
13. A method of diagnosing a serous ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein a modified level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has a serous ovarian cancer, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:
- (i) a sequence comprising at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 2 or as set forth in Table 1 and having an Accession



Number selected from the group consisting of: U62801, D49441, X51630, And AB018305;

- 5 (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 2 or as set forth in Table 1 and having an Accession Number selected from the group consisting of: U62801, D49441, X51630, And AB018305;
- (iii) a sequence that is at least about 80% identical to (i) or (ii);
- (iv) a sequence that encodes a polypeptide encoded by a nucleic acid set forth in Table 2 or as set forth in Table 1 and having an Accession Number selected from the group consisting of: U62801, D49441, X51630, And AB018305; and
- 10 (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

14. A method of diagnosing a mucinous ovarian cancer in a human or animal subject  
15 being tested said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein an elevated level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being  
20 tested has a mucinous ovarian cancer, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:

- (i) a sequence comprising at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: NM\_006149, AA315933, U47732, NM\_005588, AW503395,  
25 NM\_004063, AI073913, AI928445, NM\_022454, W40460, AA132961 and AF111856;
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of:  
30 NM\_006149, AA315933, U47732, NM\_005588, AW503395, NM\_004063, AI073913, AI928445, NM\_022454, W40460, AA132961 and AF111856;
- (iii) a sequence that is at least about 80% identical to (i) or (ii);
- (iv) a sequence that encodes a polypeptide encoded by a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of:  
35 NM\_006149, AA315933, U47732, NM\_005588, AW503395, NM\_004063, AI073913, AI928445, NM\_022454, W40460, AA132961 and AF111856; and

- (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

15      15. The method of claim 14 wherein the nucleic acid probe comprises a sequence selected from the group consisting of:

- (i) a sequence comprising at least about 20 contiguous nucleotides from SEQ ID NO: 57 or 59 or 61;
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from SEQ ID NO: 57 or 59 or 61;
- 10      (iii) a sequence that is at least about 80% identical to SEQ ID NO: 57 or 59 or 61;
- (iv) a sequence that encodes the amino acid sequence set forth in SEQ ID NO: 58 or 60 or 62; and
- (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

15

16. The method according to any one of claims 1 to 15 comprising performing a PCR reaction.

20      17. The method according to any one of claims 1 to 16 comprising performing a nucleic acid hybridization.

18. A method of detecting an ovarian cancer-associated polypeptide in a biological sample the method comprising contacting the biological sample with an antibody that binds specifically to an ovarian cancer-associated polypeptide in the biological sample, 25 the polypeptide being encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-3.

19. A method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being 30 tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein a modified level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said antibody binds to a 35 polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a sequence having at least about 80% identity to a sequence

selected from the group consisting of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 47, 49, 51, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84.

- 5 20. The method of claim 19 wherein the antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a sequence having at least about 80% identity to a sequence selected from the group consisting of SEQ ID NOs: 2, 6, 8, 10, 12, 14, 16, 18, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 47, 49, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84.

10

21. A method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein an enhanced level of the  
15 antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has an ovarian cancer, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a polypeptide encoded by a nucleic acid set forth in  
20 Table 1 or 2 other than a nucleic acid having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200.

22. The method of claim 21 wherein the antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of  
25 a sequence having at least about 80% identity to a sequence selected from the group consisting of SEQ ID NOs: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 47, 49, 51, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84.

23. A method of diagnosing an ovarian cancer in a human or animal subject being  
30 tested said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein a reduced level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that  
35 the subject being tested has an ovarian cancer, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous

amino acid residues of a polypeptide encoded by a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of NM\_022117, NM\_005460, NM\_002387, AI745249 and AI694200.

- 5 24. The method of claim 23 wherein the antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a sequence having at least about 80% identity to a sequence selected from the group consisting of SEQ ID NOs: 2, 4, and 6.
- 10 25. The method according to any one of claims 19 to 24 wherein the ovarian cancer that is diagnosed is an epithelial ovarian cancer.
26. The method according to any one of claims 19 to 24 wherein the ovarian cancer that is diagnosed is selected from the group consisting of serous ovarian cancer, non-  
15 invasive ovarian cancer, mixed phenotype ovarian cancer, mucinous ovarian cancer, endometrioid ovarian cancer, clear cell ovarian cancer, papillary serous ovarian cancer, Brenner cell and undifferentiated adenocarcinoma.
27. The method according to claim 26 wherein the ovarian cancer that is diagnosed is  
20 selected from the group consisting of serous ovarian cancer, mucinous ovarian cancer, endometrioid ovarian cancer and clear cell ovarian cancer.
28. A method of diagnosing a serous ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject  
25 being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein a modified level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has a serous ovarian cancer, and wherein said  
30 antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a polypeptide encoded by a nucleic acid set forth in Table 2 or as set forth in Table 1 and having an Accession Number selected from the group consisting of: U62801, D49441, X51630, And AB018305.
- 35 29. A method of diagnosing a mucinous ovarian cancer in a human or animal subject being tested said method comprising contacting a biological sample from said subject

being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein a reduced level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has a mucinous ovarian cancer, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a polypeptide encoded by a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: NM\_006149, AA315933, U47732, NM\_005588, AW503395, NM\_004063, AI073913, AI928445, NM\_022454, W40460, AA132961 and AF111856.

30. The method according to claim 29 wherein the antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a sequence having at least about 80% identity to SEQ ID NO: 58 or 60 or 62.

31. A method of detecting an ovarian cancer-associated antibody in a biological sample the method comprising contacting the biological sample with a polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-3, wherein the polypeptide specifically binds to the ovarian cancer-associated antibody.

32. The method according to any one of claims 1 to 31 wherein the biological sample is contacted with a plurality of nucleic acid probes and/or antibodies and/or polypeptides.

33. The method according to any one of claims 1 to 32 wherein the subject being tested is a patient undergoing a therapeutic regimen to treat ovarian cancer.

34. The method according to any one of claims 1 to 32 wherein the subject being tested is a subject suspected of having ovarian cancer.

35. A method of monitoring the efficacy of a therapeutic treatment of ovarian cancer, the method comprising:

- (i) providing a biological sample from a patient undergoing the therapeutic treatment; and
- (ii) determining the level of a ovarian cancer-associated transcript in the biological sample by contacting the biological sample with a polynucleotide

that selectively hybridizes to a sequence having at least about 80% identity to a sequence as shown in any one of Tables 1-3, thereby monitoring the efficacy of the therapy.

5 36. The method according to claim 35 further comprising comparing the level of the ovarian cancer-associated transcript to a level of the ovarian cancer-associated transcript in a biological sample from the patient prior to, or earlier in, the therapeutic treatment.

10 37. A method of monitoring the efficacy of a therapeutic treatment of ovarian cancer, the method comprising :

- (i) providing a biological sample from a patient undergoing the therapeutic treatment; and
- (ii) determining the level of a ovarian cancer-associated antibody in the biological sample by contacting the biological sample with a polypeptide  
15 encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-3, wherein the polypeptide specifically binds to the ovarian cancer-associated antibody, thereby monitoring the efficacy of the therapy.

20 38. The method of claim 37 further comprising comparing the level of the ovarian cancer-associated antibody to a level of the ovarian cancer-associated antibody in a biological sample from the patient prior to, or earlier in, the therapeutic treatment.

25 39. A method of monitoring the efficacy of a therapeutic treatment of ovarian cancer, the method comprising :

- (i) providing a biological sample from a patient undergoing the therapeutic treatment; and
- (ii) determining the level of a ovarian cancer-associated polypeptide in the biological sample by contacting the biological sample with an antibody, wherein  
30 the antibody specifically binds to a polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-3, thereby monitoring the efficacy of the therapy.

40. The method of claim 39 further comprising comparing the level of the ovarian cancer-associated polypeptide to a level of the ovarian cancer-associated polypeptide in a biological sample from the patient prior to, or earlier in, the therapeutic treatment.

5 41. The method according to any one of claims 35 to 40 wherein the ovarian cancer that is diagnosed is an epithelial ovarian cancer.

42. The method according to any one of claims 35 to 41 wherein the ovarian cancer that is diagnosed is selected from the group consisting of serous ovarian cancer, non-  
10 invasive ovarian cancer, mixed phenotype ovarian cancer, mucinous ovarian cancer, endometrioid ovarian cancer, clear cell ovarian cancer, papillary serous ovarian cancer, Brenner cell and undifferentiated adenocarcinoma.

43. The method according to claim 42 wherein the ovarian cancer that is diagnosed is  
15 selected from the group consisting of serous ovarian cancer, mucinous ovarian cancer, endometrioid ovarian cancer and clear cell ovarian cancer.

44. A method of determining the likelihood of survival of a subject suffering from an ovarian cancer, said method comprising contacting a biological sample from said subject  
20 being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein an elevated level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has a poor probability of survival, and wherein said nucleic acid probe comprises a  
25 sequence selected from the group consisting of:

(i) a sequence comprising at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: NM\_003014, AA046217, NM\_015902, T83882, AB040888, AA628980, AI623351, AW614420, AA243499, AF251237, AI970797, AF145713,  
30 X78565, T97307, BE243845, AW068302, AL133561, BE313555, X07820, AI973016, AF084545, U41518, Z11894, AW138190, BE086548, W47196, AI796870, X02761, AW968613, AW972565, AF045229, AW953853, U52426, F06700, AI798863, H52761, BE546947, AU076643, U20536, AA581602, AJ245210, X65965, AI806770, BE386490, AW581992, U77534, AL034417,  
35 L10343, AW518944, W28729, AI640160, U11862, AW295980, X59135, BE466173, AI354722, M90464, AA829286, AI333771, BE465867, NM\_014992,

- BE616902, AA430373, R27430, BE387335, AW264102, AW952323, AA088177, BE614567, AL079658, NM\_002776, BE261944, NM\_006379, AI002238, X81789, NM\_002122, AB001914, AA311919, AI381750, AA292998, BE439580, AI677897, N72403, BE003054, AL035588, AI080491, AW770994, H24177, AF146761, NM\_001955, AI680737, AI752666, AA505445, BE246649, and NM\_003955;
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: NM\_003014, AA046217, NM\_015902, T83882, AB040888, AA628980, AI623351, AW614420, AA243499, AF251237, AI970797, AF145713, X78565, T97307, BE243845, AW068302, AL133561, BE313555, X07820, AI973016, AF084545, U41518, Z11894, AW138190, BE086548, W47196, AI796870, X02761, AW968613, AW972565, AF045229, AW953853, U52426, F06700, AI798863, H52761, BE546947, AU076643, U20536, AA581602, AJ245210, X65965, AI806770, BE386490, AW581992, U77534, AL034417, L10343, AW518944, W28729, AI640160, U11862, AW295980, X59135, BE466173, AI354722, M90464, AA829286, AI333771, BE465867, NM\_014992, BE616902, AA430373, R27430, BE387335, AW264102, AW952323, AA088177, BE614567, AL079658, NM\_002776, BE261944, NM\_006379, AI002238, X81789, NM\_002122, AB001914, AA311919, AI381750, AA292998, BE439580, AI677897, N72403, BE003054, AL035588, AI080491, AW770994, H24177, AF146761, NM\_001955, AI680737, AI752666, AA505445, BE246649, and NM\_003955;
- (iii) a sequence that is at least about 80% identical to (i) or (ii);
- (iv) a sequence that encodes a polypeptide encoded by a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: NM\_003014, AA046217, NM\_015902, T83882, AB040888, AA628980, AI623351, AW614420, AA243499, AF251237, AI970797, AF145713, X78565, T97307, BE243845, AW068302, AL133561, BE313555, X07820, AI973016, AF084545, U41518, Z11894, AW138190, BE086548, W47196, AI796870, X02761, AW968613, AW972565, AF045229, AW953853, U52426, F06700, AI798863, H52761, BE546947, AU076643, U20536, AA581602, AJ245210, X65965, AI806770, BE386490, AW581992, U77534, AL034417, L10343, AW518944, W28729, AI640160, U11862, AW295980, X59135, BE466173, AI354722, M90464, AA829286, AI333771, BE465867, NM\_014992, BE616902, AA430373, R27430, BE387335, AW264102, AW952323, AA088177, BE614567, AL079658, NM\_002776, BE261944, NM\_006379, AI002238, X81789, NM\_002122,



AB001914, AA311919, AI381750, AA292998, BE439580, AI677897, N72403, BE003054, AL035588, AI080491, AW770994, H24177, AF146761, NM\_001955, AI680737, AI752666, AA505445, BE246649, and NM\_003955; and

- 5 (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

45. The method of claim 44 wherein the nucleic acid probe comprises a sequence selected from the group consisting of:

- 10 (i) a sequence comprising at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 63, 65, 67, 69, 71, and 73;
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- 15 (iii) a sequence that is at least about 80% identical to a sequence selected from the group consisting of SEQ ID NOs: 63, 65, 67, 69, 71, 73, 75, 77, 79, 81 and 83;
- (iv) a sequence that encodes an amino acid sequence selected from the group consisting of SEQ ID NOs: 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84; and
- (v) a sequence that is complementary to (i) or (ii) or (iii) or (iv).

- 20 46. A method of determining the likelihood of survival of a subject suffering from an ovarian cancer, said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein an enhanced level of the antigen-antibody complex for the subject being tested compared to the amount of the
- 25 antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has a poor probability of survival, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a sequence encoded by a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting
- 30 of: NM\_003014, AA046217, NM\_015902, T83882, AB040888, AA628980, AI623351, AW614420, AA243499, AF251237, AI970797, AF145713, X78565, T97307, BE243845, AW068302, AL133561, BE313555, X07820, AI973016, AF084545, U41518, Z11894, AW138190, BE086548, W47196, AI796870, X02761, AW968613, AW972565, AF045229, AW953853, U52426, F06700, AI798863, H52761, BE546947, AU076643,
- 35 U20536, AA581602, AJ245210, X65965, AI806770, BE386490, AW581992, U77534, AL034417, L10343, AW518944, W28729, AI640160, U11862, AW295980, X59135,

BE466173, AI354722, M90464, AA829286, AI333771, BE465867, NM\_014992, BE616902, AA430373, R27430, BE387335, AW264102, AW952323, AA088177, BE614567, AL079658, NM\_002776, BE261944, NM\_006379, AI002238, X81789, NM\_002122, AB001914, AA311919, AI381750, AA292998, BE439580, AI677897, 5 N72403, BE003054, AL035588, AI080491, AW770994, H24177, AF146761, NM\_001955, AI680737, AI752666, AA505445, BE246649, and NM\_003955.

47. The method of claim 46 wherein the antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of 10 a sequence having at least about 80% identity to a sequence selected from the group consisting of SEQ ID NOs: 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 and 84.

48. A method of determining the likelihood of survival of a subject suffering from a serous ovarian cancer, said method comprising contacting a biological sample from said 15 subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein an elevated level of hybridization of the probe for the subject being tested compared to the hybridization obtained for a control subject not having ovarian cancer indicates that the subject being tested has a poor probability of survival, and wherein said nucleic acid probe comprises a 20 sequence selected from the group consisting of:

- (i) a sequence comprising at least about 20 contiguous nucleotides from a nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO: 71 or 73;
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a nucleic acid comprising the 25 nucleotide sequence set forth in SEQ ID NO: 71 or 73;
- (iii) a sequence that is at least about 80% identical to (i) or (ii) and encoding an sFRP protein or a SOCS3 protein;
- (iv) a sequence that encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 72 or 74; and
- 30 (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

49. A method of determining the likelihood of survival of a subject suffering from a serous ovarian cancer, said method comprising contacting a biological sample from said 35 subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein an enhanced

level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has a poor probability of survival, and wherein said antibody binds to an sFRP polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 72 or a SOCS3 polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 74.

50. A method of determining the likelihood of survival of a subject suffering from a serous ovarian cancer, said method comprising contacting a biological sample from said subject being tested with at least two antibodies for a time and under conditions sufficient for antigen-antibody complexes to form and then detecting the complexes wherein an enhanced level of the antigen-antibody complexes for the subject being tested compared to the amount of the antigen-antibody complexes formed for a control subject not having ovarian cancer indicates that the subject being tested has a poor probability of survival, and wherein one antibody binds to an sFRP polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 72 and wherein one antibody binds to a SOCS3 polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 74.

51. The method according to any one of claims 44 to 47 wherein the ovarian cancer is an epithelial ovarian cancer.

52. The method according to any one of claims 44 to 47 wherein the ovarian cancer is selected from the group consisting of serous ovarian cancer, non-invasive ovarian cancer, mixed phenotype ovarian cancer, mucinous ovarian cancer, endometrioid ovarian cancer, clear cell ovarian cancer, papillary serous ovarian cancer, Brenner cell and undifferentiated adenocarcinoma.

53. The method according to claim 52 wherein the ovarian cancer is selected from the group consisting of serous ovarian cancer, mucinous ovarian cancer, endometrioid ovarian cancer and clear cell ovarian cancer.

54. A method of determining the likelihood that a subject will suffer from a recurrence of an ovarian cancer, said method comprising contacting a biological sample from said subject being tested with a nucleic acid probe for a time and under conditions sufficient for hybridization to occur and then detecting the hybridization wherein an elevated level of hybridization of the probe for the subject being tested compared to the hybridization

obtained for a control subject not having ovarian cancer indicates that the subject being tested has a high probability of recurrence, and wherein said nucleic acid probe comprises a sequence selected from the group consisting of:

- (i) a sequence comprising at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: M86849, AW963419, BE298665, AK000637, BE077546, T97307, R24601, BE090176, AA393907, W28729, BE313754, AW673081, AA356694, L08239, BE397649, NM\_012317, NM\_000947, AJ250562, AL040183, BE207573, BE564162, BE439580, AW067800, AA569756, AW138190, AF126245, L10343, NM\_002514, AI863735, NM\_005397, W26391, H15474, U51166, AA243499, AW408807, AI738719, AB040888, BE313077, AI677897, C14898, AI821730, AF007393, H65423, N46243, AA095971, U20350, NM\_005756, D19589, AW957446, AW294647, BE159718, AI888490, AA022569, BE147740, AI798863, BE464341, AL080235, AI557212, X75208, AA628980, BE242587, NM\_005512, AW953853, AU076611, AW968613, AL353944, BE614149, AA292998, H12912, AA188763, AK000596, AI970797, AW519204, Z42387, AF145713, AA972412, AK001564, AW959861, BE313555, W25005, AI193356, AF111106, AI130740, AA985190, BE221880, AF084545, R26584, AW247380, AA364261, U25849, AF262992, AW342140, AL133572, AI497778, AI745379, U51712, AW375974, AF251237, NM\_000636, AA130986, AA216363, AA628980, AA811657, AA897108, AB040888, AF212225, AI089575, AI282028, AI368826, AI718702, AI827248, AK002039, AL109791, AW090198, AW296454, AW445034, AW452948, AW470411, AW885727, AW970859, AW979189, BE165866, BE175582, BE242587, BE271927, BE439580, BE464016, D63216, F34856, M83822, N33937, N49068, N51357, N80486, NM\_000954, NM\_005756, NM\_016652, R26584, R31178, W05391, W25005, W45393, W68815, X65965, X76732 and Z45051,
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to at least about 20 contiguous nucleotides from a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: M86849, AW963419, BE298665, AK000637, BE077546, T97307, R24601, BE090176, AA393907, W28729, BE313754, AW673081, AA356694, L08239, BE397649, NM\_012317, NM\_000947, AJ250562, AL040183, BE207573, BE564162, BE439580, AW067800, AA569756, AW138190, AF126245, L10343, NM\_002514, AI863735, NM\_005397, W26391, H15474, U51166, AA243499, AW408807, AI738719, AB040888, BE313077, AI677897, C14898, AI821730, AF007393,

H65423, N46243, AA095971, U20350, NM\_005756, D19589, AW957446,  
 AW294647, BE159718, AI888490, AA022569, BE147740, AI798863, BE464341,  
 AL080235, AI557212, X75208, AA628980, BE242587, NM\_005512, AW953853,  
 AU076611, AW968613, AL353944, BE614149, AA292998, H12912, AA188763,  
 5 AK000596, AI970797, AW519204, Z42387, AF145713, AA972412, AK001564,  
 AW959861, BE313555, W25005, AI193356, AF111106, AI130740, AA985190,  
 BE221880, AF084545, R26584, AW247380, AA364261, U25849, AF262992,  
 AW342140, AL133572, AI497778, AI745379, U51712, AW375974, AF251237,  
 NM\_000636, AA130986, AA216363, AA628980, AA811657, AA897108,  
 10 AB040888, AF212225, AI089575, AI282028, AI368826, AI718702, AI827248,  
 AK002039, AL109791, AW090198, AW296454, AW445034, AW452948,  
 AW470411, AW885727, AW970859, AW979189, BE165866, BE175582,  
 BE242587, BE271927, BE439580, BE464016, D63216, F34856, M83822,  
 N33937, N49068, N51357, N80486, NM\_000954, NM\_005756, NM\_016652,  
 15 R26584, R31178, W05391, W25005, W45393, W68815, X65965, X76732 and  
 Z45051;

(iii) a sequence that is at least about 80% identical to (i) or (ii);

(iv) a sequence that encodes a polypeptide encoded by a nucleic acid set forth in  
 Table 1 and having an Accession Number selected from the group consisting  
 20 of: M86849, AW963419, BE298665, AK000637, BE077546, T97307, R24601,  
 BE090176, AA393907, W28729, BE313754, AW673081, AA356694, L08239,  
 BE397649, NM\_012317, NM\_000947, AJ250562, AL040183, BE207573,  
 BE564162, BE439580, AW067800, AA569756, AW138190, AF126245, L10343,  
 NM\_002514, AI863735, NM\_005397, W26391, H15474, U51166, AA243499,  
 25 AW408807, AI738719, AB040888, BE313077, AI677897, C14898, AI821730,  
 AF007393, H65423, N46243, AA095971, U20350, NM\_005756, D19589,  
 AW957446, AW294647, BE159718, AI888490, AA022569, BE147740, AI798863,  
 BE464341, AL080235, AI557212, X75208, AA628980, BE242587, NM\_005512,  
 AW953853, AU076611, AW968613, AL353944, BE614149, AA292998, H12912,  
 30 AA188763, AK000596, AI970797, AW519204, Z42387, AF145713, AA972412,  
 AK001564, AW959861, BE313555, W25005, AI193356, AF111106, AI130740,  
 AA985190, BE221880, AF084545, R26584, AW247380, AA364261, U25849,  
 AF262992, AW342140, AL133572, AI497778, AI745379, U51712, AW375974,  
 AF251237, NM\_000636, AA130986, AA216363, AA628980, AA811657,  
 35 AA897108, AB040888, AF212225, AI089575, AI282028, AI368826, AI718702,  
 AI827248, AK002039, AL109791, AW090198, AW296454, AW445034,

AW452948, AW470411, AW885727, AW970859, AW979189, BE165866, BE175582, BE242587, BE271927, BE439580, BE464016, D63216, F34856, M83822, N33937, N49068, N51357, N80486, NM\_000954, NM\_005756, NM\_016652, R26584, R31178, W05391, W25005, W45393, W68815, X65965, X76732 and Z45051; and

- (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

55. The method of claim 54 determining the likelihood that a subject will suffer from a recurrence of an ovarian cancer, said method comprising contacting a biological sample from said subject being tested with an antibody for a time and under conditions sufficient for an antigen-antibody complex to form and then detecting the complex wherein an enhanced level of the antigen-antibody complex for the subject being tested compared to the amount of the antigen-antibody complex formed for a control subject not having ovarian cancer indicates that the subject being tested has a high probability of recurrence, and wherein said antibody binds to a polypeptide comprising an amino acid sequence comprising at least about 10 contiguous amino acid residues of a sequence encoded by a nucleic acid set forth in Table 1 and having an Accession Number selected from the group consisting of: M86849, AW963419, BE298665, AK000637, BE077546, T97307, R24601, BE090176, AA393907, W28729, BE313754, AW673081, AA356694, L08239, BE397649, NM\_012317, NM\_000947, AJ250562, AL040183, BE207573, BE564162, BE439580, AW067800, AA569756, AW138190, AF126245, L10343, NM\_002514, AI863735, NM\_005397, W26391, H15474, U51166, AA243499, AW408807, AI738719, AB040888, BE313077, AI677897, C14898, AI821730, AF007393, H65423, N46243, AA095971, U20350, NM\_005756, D19589, AW957446, AW294647, BE159718, AI888490, AA022569, BE147740, AI798863, BE464341, AL080235, AI557212, X75208, AA628980, BE242587, NM\_005512, AW953853, AU076611, AW968613, AL353944, BE614149, AA292998, H12912, AA188763, AK000596, AI970797, AW519204, Z42387, AF145713, AA972412, AK001564, AW959861, BE313555, W25005, AI193356, AF111106, AI130740, AA985190, BE221880, AF084545, R26584, AW247380, AA364261, U25849, AF262992, AW342140, AL133572, AI497778, AI745379, U51712, AW375974, AF251237, NM\_000636, AA130986, AA216363, AA628980, AA811657, AA897108, AB040888, AF212225, AI089575, AI282028, AI368826, AI718702, AI827248, AK002039, AL109791, AW090198, AW296454, AW445034, AW452948, AW470411, AW885727, AW970859, AW979189, BE165866, BE175582, BE242587, BE271927, BE439580,

BE464016, D63216, F34856, M83822, N33937, N49068, N51357, N80486, NM\_000954, NM\_005756, NM\_016652, R26584, R31178, W05391, W25005, W45393, W68815, X65965, X76732 and Z45051.

- 5 56. The method according to claim 54 or 55 wherein the ovarian cancer is an epithelial ovarian cancer.
57. The method according to any one of claims 54 to 56 wherein the ovarian cancer is selected from the group consisting of serous ovarian cancer, non-invasive ovarian  
10 cancer, mixed phenotype ovarian cancer, mucinous ovarian cancer, endometrioid ovarian cancer, clear cell ovarian cancer, papillary serous ovarian cancer, Brenner cell and undifferentiated adenocarcinoma.
58. The method according to claim 57 wherein the ovarian cancer is selected from the  
15 group consisting of serous ovarian cancer, mucinous ovarian cancer, endometrioid ovarian cancer and clear cell ovarian cancer.
59. The method according to any one of claims 35 to 58 wherein the biological sample is contacted with a plurality of nucleic acid probes and/or antibodies and/or polypeptides.  
20
60. A method for identifying a compound that modulates an ovarian cancer-associated polypeptide, the method comprising :
- 25 (i) contacting the compound with a ovarian cancer-associated polypeptide, the polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-3; and
- (ii) determining the functional effect of the compound upon the polypeptide.
61. A method for determining a candidate compound for the treatment of ovarian cancer comprising :
- 30 (i) administering a test compound to a mammal having ovarian cancer or a cell isolated therefrom;
- (ii) comparing the level of gene expression of a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-3 in a treated cell or mammal with the level of gene expression of the  
35 polynucleotide in a control cell or mammal, wherein a test compound that

modulates the level of expression of the polynucleotide is a candidate for the treatment of ovarian cancer.

- 5 62. An assay device for use in the diagnosis or prognosis of ovarian cancer, said device comprising a plurality of polynucleotides immobilized to a solid phase, wherein each of said polynucleotides consists of a gene as listed in any one of Tables 1-3.
63. The device of claim 62 consisting of a substantially planar chip.
- 10 64. An assay device for use in the diagnosis or prognosis of ovarian cancer, said device comprising a plurality of different antibodies immobilized to a solid phase, wherein each of said antibodies binds to a polypeptide listed in Tables 1-3.
65. The device of claim 64 consisting of a substantially planar chip.
- 15 66. Use of a polynucleotide as set forth in any one of Tables 1-3 in the diagnosis or prognosis of ovarian cancer or for the preparation of a medicament for the treatment of ovarian cancer.
- 20 67. Use of a vector comprising a polynucleotide as set forth in any one of Tables 1-3 in the diagnosis or prognosis of ovarian cancer or for the preparation of a medicament for the treatment of ovarian cancer.
- 25 68. Use of an isolated polypeptide as set forth in any one of Tables 1-3 in the diagnosis or prognosis of ovarian cancer or for the preparation of a medicament for the treatment of ovarian cancer.
- 30 69. Use of an antibody that binds to an isolated polypeptide as set forth in any one of Tables 1-3 in the diagnosis or prognosis of ovarian cancer or for the preparation of a medicament for the treatment of ovarian cancer.
70. A method of diagnosing an ovarian cancer in a human or animal subject being tested said method comprising determining aberrant methylation in a promoter sequence that regulates expression of a tumor suppressor gene in a biological sample from said  
35 subject compared to the methylation of the promoter in nucleic acid obtained for a control



subject not having ovarian cancer wherein said aberrant methylation indicates that the subject being tested has an ovarian ovarian cancer.

5 71. The method of claim 70 wherein hypermethylation of the promoter sequence is determined.

72. The method of claim 70 or 71 wherein the methylation is determined in the promoter region that regulates expression of an MCC gene comprising a sequence selected from the group consisting of:

- 10 (i) the nucleotide sequence set forth as SEQ ID NO: 3;
- (ii) a sequence that hybridizes under at least low stringency hybridization conditions to the nucleotide sequence set forth as SEQ ID NO: 3;
- (iii) a sequence that is at least about 80% identical to (i) or (ii);
- 15 (iv) a sequence that encodes a polypeptide encoded by a nucleotide sequence set forth as SEQ ID NO: 3; and
- (v) a sequence that is complementary to any one of the sequences set forth in (i) or (ii) or (iii) or (iv).

20 73. The method according to any one of claims 70 to 72 wherein the ovarian cancer that is diagnosed is an epithelial ovarian cancer.

1/5

Claudin 3

EP-CAM

SOX17

CA125

MUC-1

E-cadherin

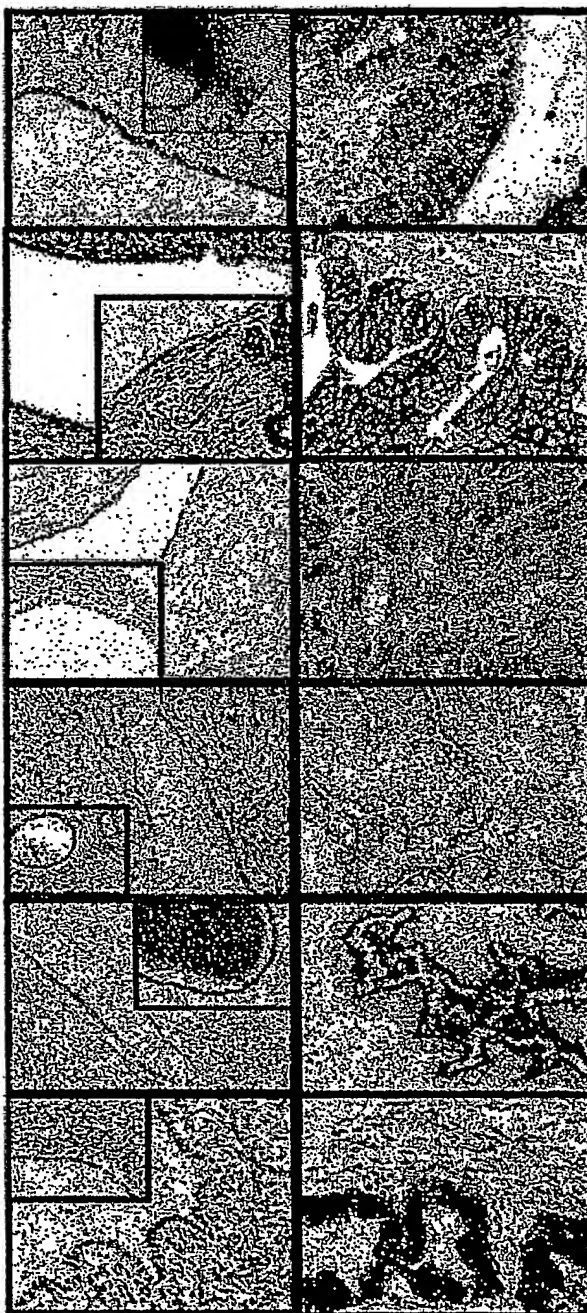


FIGURE 1

626929/01

2/5

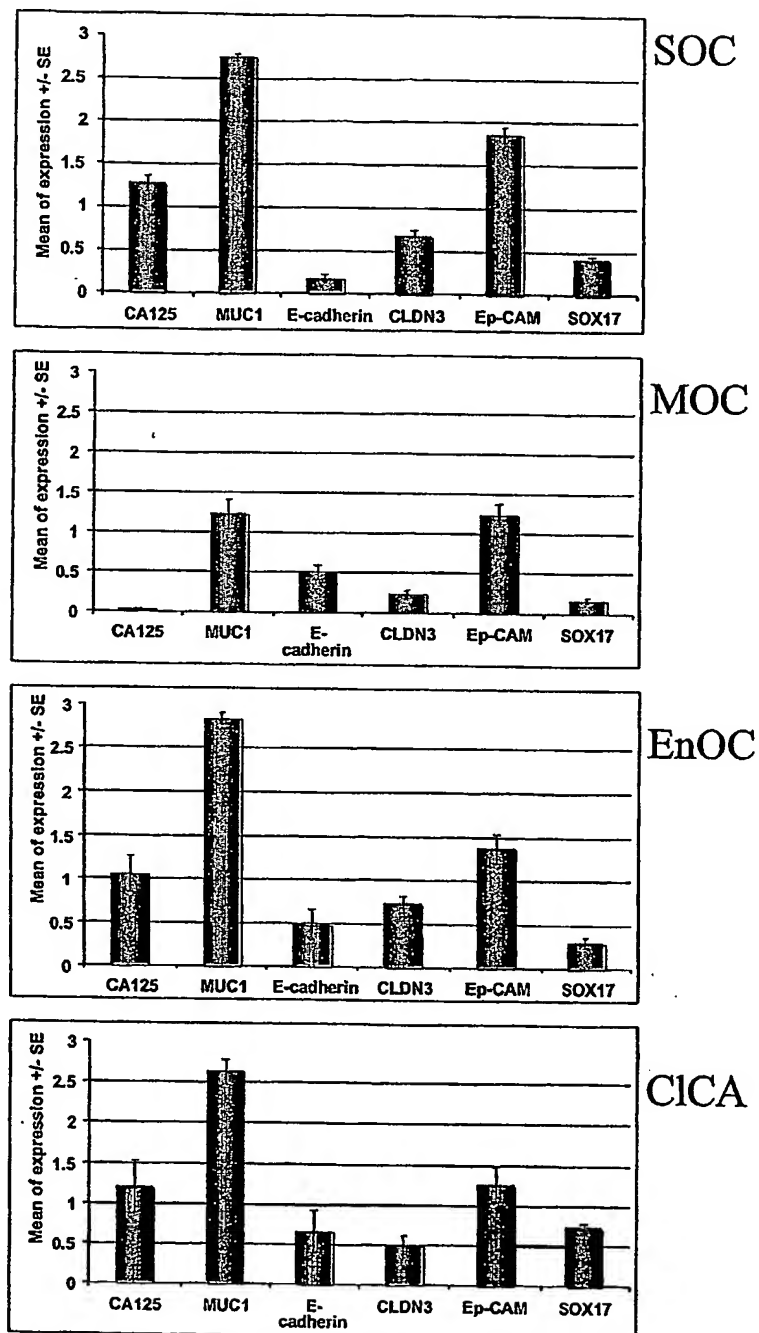


FIGURE 2

626975/01

3/5

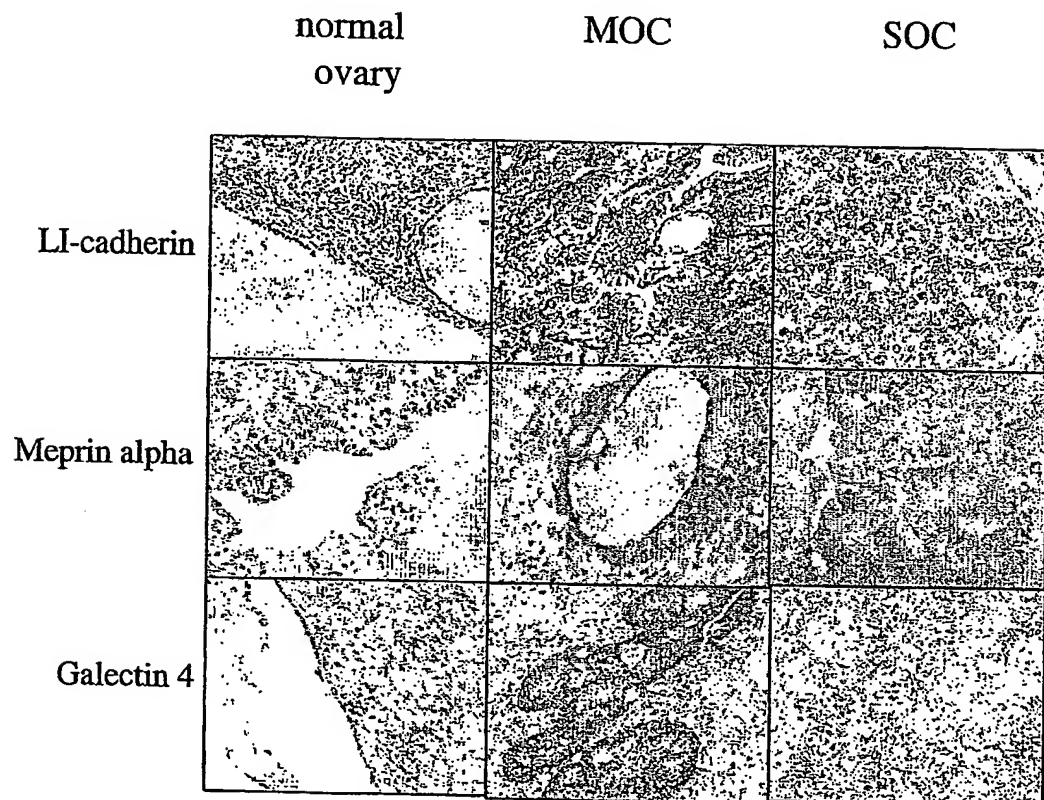


FIGURE 3

10/526979

4/5

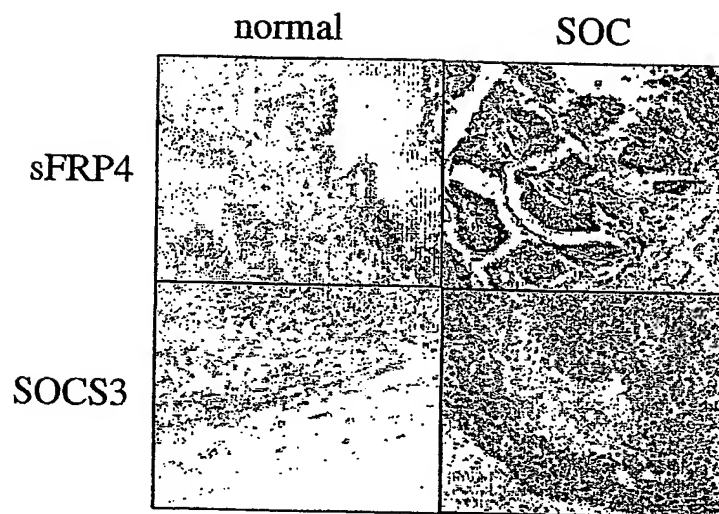


FIGURE 4a

626929/01

5/5

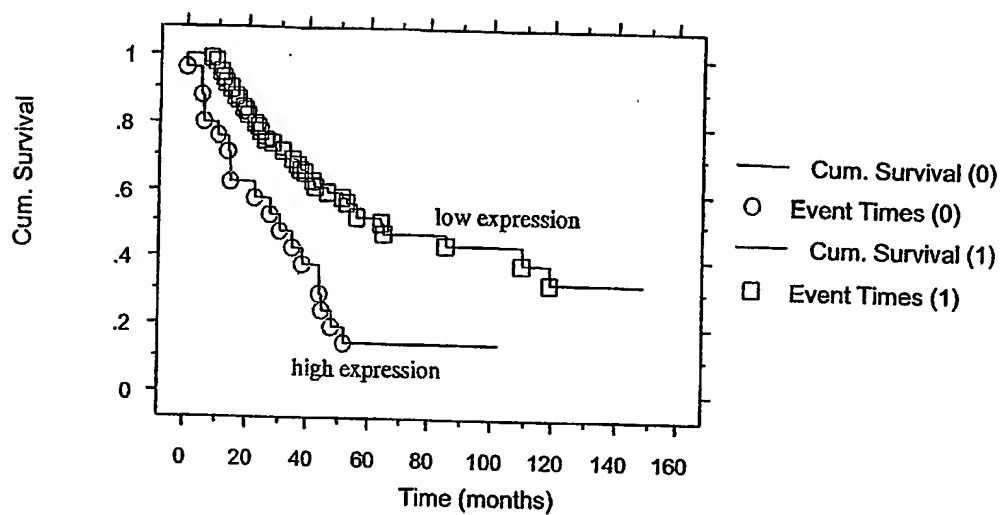


FIGURE 4b

10/526979

01013801010

01013801010

DT01 Rec'd PCT/PTO 07 MAR 2005

## SEQUENCE LISTING

<110> Garvan Institute of Medical Research

<120> Methods for the diagnosis and prognosis of ovarian cancer

<130> 501731/MRO

<150> AU2002951346

<151> 2002-09-05

<160> 84

<170> PatentIn version 3.1

<210> 1

<211> 3432

<212> DNA

<213> NM\_005460 SNCAIP, synphilin

<220>

<221> CDS

<222> (94)..(2850)

<223>

<400> 1

gtcagtcctcct tcgcgctcct gagccgccgg cgccgccggg gccccgggaa ttataagta 60

tttgaccgta ctcaaatgt gcaaggaaga ata atg gaa gcc cct gaa tac ctt 114

Met Glu Ala Pro Glu Tyr Leu

1 5

gat ttg gat gaa att gac ttt agt gat gac ata tct tat tca gtc aca 162

Asp Leu Asp Glu Ile Asp Phe Ser Asp Asp Ile Ser Tyr Ser Val Thr

10 15 20

tca ctc aag acg atc cca gaa ctg tgc cga aga tgt gat acg caa aac 210

Ser Leu Lys Thr Ile Pro Glu Leu Cys Arg Arg Cys Asp Thr Gln Asn

25 30 35

gaa gac aga tca gct tct agc tct agc tgg aat tgt ggc atc tca act 258

Glu Asp Arg Ser Ala Ser Ser Ser Ser Trp Asn Cys Gly Ile Ser Thr

40 45 50 55

ctt att aca aac acg caa aag ccc aca gga atc gct gat gtg tac agt 306

Leu Ile Thr Asn Thr Gln Lys Pro Thr Gly Ile Ala Asp Val Tyr Ser

60 65 70

aag ttc cgc cca gtg aag cgg gtt tgc cca ctg aaa cat cag cca gag 354

Lys Phe Arg Pro Val Lys Arg Val Ser Pro Leu Lys His Gln Pro Glu

75 80 85

act ctg gag aac aat gaa agt gat gac caa aag aac cag aaa gtg gtt 402

Thr Leu Glu Asn Asn Glu Ser Asp Asp Gln Lys Asn Gln Lys Val Val

90 95 100

gag tac cag aaa ggg ggt gag tct gac ctg ggc ccc cag cct cag gag 450

Glu Tyr Gln Lys Gly Gly Ser Asp Leu Gly Pro Gln Pro Gln Glu

105 110 115

ctt ggc cct gga gat gga gtg ggc ggc cca cca ggt aag agc tct gag 498

Leu Gly Pro Gly Asp Gly Val Gly Gly Pro Pro Gly Lys Ser Ser Glu

120 125 130 135

ccc agc aca tcg ctg ggt gaa ctg gag cac tac gac ctc gac atg gat 546

Pro Ser Thr Ser Leu Gly Glu Leu Glu His Tyr Asp Leu Asp Met Asp

140 145 150

gag att ctg gat gtg cct tat att aaa tcc agt cag cag ctt gcc tct Glu Ile Leu Asp Val Pro Tyr Ile Lys Ser Ser Gln Gln Leu Ala Ser 155 160 165	594
ttt acc aag gtg act tca gaa aaa aga att ttg ggc tta tgc aca acc Phe Thr Lys Val Thr Ser Glu Lys Arg Ile Leu Gly Leu Cys Thr Thr 170 175 180	642
atc aat ggc ctt tct ggc aaa gcc tgc tct aca gga agt tct gag agc Ile Asn Gly Leu Ser Gly Lys Ala Cys Ser Thr Gly Ser Ser Glu Ser 185 190 195	690
tca tca tcc aac atg gca cca ttt tgt gtt ctt tct ccc gtg aaa agc Ser Ser Ser Asn Met Ala Pro Phe Cys Val Leu Ser Pro Val Lys Ser 200 205 210 215	738
cct cac ttg aga aaa gca tca gct gtc atc cac gac cag cac aag ctg Pro His Leu Arg Lys Ala Ser Ala Val Ile His Asp Gln His Lys Leu 220 225 230	786
tcc act gaa gaa acc gag atc tca cct cct ctg gtt aaa tgt ggc tct Ser Thr Glu Glu Thr Glu Ile Ser Pro Pro Leu Val Lys Cys Gly Ser 235 240 245	834
gca tat gag cct gaa aac cag agt aaa gac ttc cta aac aag aca ttt Ala Tyr Glu Pro Glu Asn Gln Ser Lys Asp Phe Leu Asn Lys Thr Phe 250 255 260	882
agt gat cct cat ggt cga aaa gtt gag aag aca aca cca gac tgc cag Ser Asp Pro His Gly Arg Lys Val Glu Lys Thr Thr Pro Asp Cys Gln 265 270 275	930
ctc agg gcc ttc cac cta caa tcc tca gca gca gaa tcc aaa cca gaa Leu Arg Ala Phe His Leu Gln Ser Ser Ala Ala Glu Ser Lys Pro Glu 280 285 290 295	978
gag cag gtc agt ggc cta aac cgg acc agc tcc caa ggc cca gaa gaa Glu Gln Val Ser Gly Leu Asn Arg Thr Ser Ser Gln Gly Pro Glu Glu 300 305 310	1026
agg agt gag tat ctg aaa aaa gtg aaa agc atc ttg aac att gtt aaa Arg Ser Glu Tyr Leu Lys Lys Val Lys Ser Ile Leu Asn Ile Val Lys 315 320 325	1074
gaa gga cag atc tct ctc ctg cca cac cta gct gca gac aat cta gac Glu Gly Gln Ile Ser Leu Leu Pro His Leu Ala Ala Asp Asn Leu Asp 330 335 340	1122
aaa att cac gac gaa aat gga aac aat cta tta cat att gcg gcg tca Lys Ile His Asp Glu Asn Gly Asn Asn Leu Leu His Ile Ala Ala Ser 345 350 355	1170
cag gga cac gca gag tgt cta cag cac ctc act tct ttg atg gga gaa Gln Gly His Ala Glu Cys Leu Gln His Leu Thr Ser Leu Met Gly Glu 360 365 370 375	1218
gac tgc ctc aat gag cgc aac act gag aag ttg act cca gca ggc ctg Asp Cys Leu Asn Glu Arg Asn Thr Glu Lys Leu Thr Pro Ala Gly Leu 380 385 390	1266
gcc att aag aat ggt cag ttg gag tgc gta cgc tgg atg gtg agc gaa Ala Ile Lys Asn Gly Gln Leu Glu Cys Val Arg Trp Met Val Ser Glu 395 400 405	1314
aca gaa gcc att gca gaa ctg agt tgt tct aag gat ttt cca agc ctt Thr Glu Ala Ile Ala Glu Leu Ser Cys Ser Lys Asp Phe Pro Ser Leu 410 415 420	1362



att cat tac gca ggt tgc tat ggc cag gaa aag att ctt ctg tgg ctt Ile His Tyr Ala Gly Cys Tyr Gly Gln Glu Lys Ile Leu Leu Trp Leu 425 430 435	1410
ctt cag ttt atg caa gaa cag ggc atc tcg ttg gat gaa gta gac cag Leu Gln Phe Met Gln Glu Gln Gly Ile Ser Leu Asp Glu Val Asp Gln 440 445 450 455	1458
gat ggc aac agt gcc gtt cac gta gcc tca cag cat ggc tac ctt gga Asp Gly Asn Ser Ala Val His Val Ala Ser Gln His Gly Tyr Leu Gly 460 465 470	1506
tgc ata cag acc ttg gtt gaa tat gga gca aat gtc acc atg cag aac Cys Ile Gln Thr Leu Val Glu Tyr Gly Ala Asn Val Thr Met Gln Asn 475 480 485	1554
cac gct ggg gaa aag ccc tcc cag agc gcc gag cgg cag ggg cac acc His Ala Gly Glu Lys Pro Ser Gln Ser Ala Glu Arg Gln Gly His Thr 490 495 500	1602
ctg tgc tcc agg tac ctg gtg gtg gtg gag acc tgc atg tcg ctg gcc Leu Cys Ser Arg Tyr Leu Val Val Val Glu Thr Cys Met Ser Leu Ala 505 510 515	1650
tct caa gtg gtg aag tta acc aag cag cta aag gaa caa aca gta gaa Ser Gln Val Val Lys Leu Thr Lys Gln Leu Lys Glu Gln Thr Val Glu 520 525 530 535	1698
cgt gtc acg ctg cag aac caa ctc caa caa ttt cta gaa gcc cag aaa Arg Val Thr Leu Gln Asn Gln Leu Gln Phe Leu Glu Ala Gln Lys 540 545 550	1746
tca gag ggc aag tca ctc cct tct tca ccc agt tca cca tcc tca cct Ser Glu Gly Lys Ser Leu Pro Ser Ser Pro Ser Ser Pro Ser Pro 555 560 565	1794
gcc tcc aga aag tcc cag tgg aaa tct cca gat gca gat gat gat tct Ala Ser Arg Lys Ser Gln Trp Lys Ser Pro Asp Ala Asp Asp Ser 570 575 580	1842
gta gcc aaa agc aag cca gga gtc caa gag ggg att cag gtt ctt gga Val Ala Lys Ser Lys Pro Gly Val Gln Glu Gly Ile Gln Val Leu Gly 585 590 595	1890
agc ctg tca gcc tcc agc cgg gct aga ccc aaa gca aaa gat gaa gat Ser Leu Ser Ala Ser Ser Arg Ala Arg Pro Lys Ala Lys Asp Glu Asp 600 605 610 615	1938
tct gat aaa atc tta cgc cag tta ttg gga aag gaa atc tca gaa aat Ser Asp Lys Ile Leu Arg Gln Leu Leu Gly Lys Glu Ile Ser Glu Asn 620 625 630	1986
gtc tgc acc cag gaa aaa ctg tcc ttg gaa ttc cag gat gct cag gct Val Cys Thr Gln Glu Lys Leu Ser Leu Glu Phe Gln Asp Ala Gln Ala 635 640 645	2034
tcc tct aga aat tct aaa aag atc cca ctg gag aag agg gaa ctg aag Ser Ser Arg Asn Ser Lys Lys Ile Pro Leu Glu Lys Arg Glu Leu Lys 650 655 660	2082
tta gcc agg ctg aga cag ctg atg cag agg tca ctg agt gag tct gac Leu Ala Arg Leu Arg Gln Leu Met Gln Arg Ser Leu Ser Glu Ser Asp 665 670 675	2130
aca gac tcc aac aac tct gag gac ccc aag act acc cca gtg agg aag Thr Asp Ser Asn Asn Ser Glu Asp Pro Lys Thr Thr Pro Val Arg Lys 680 685 690 695	2178

gct gac cga cca agg ccg cag ccc att gta gaa agc gta gag agt atg Ala Asp Arg Pro Arg Pro Gln Pro Ile Val Glu Ser Val Glu Ser Met 700 705 710	2226
gac agc gca gaa agc ctg cac ctg atg att aag aaa cac acc ttg gca Asp Ser Ala Glu Ser Leu His Leu Met Ile Lys Lys His Thr Leu Ala 715 720 725	2274
tca ggg gga cgc agg ttt cct ttc agc atc aag gcc tcc aaa tcc ctg Ser Gly Gly Arg Arg Phe Pro Phe Ser Ile Lys Ala Ser Lys Ser Leu 730 735 740	2322
gat ggc cac agc cca tct ccc acc tca gag agc agc gaa cca gac tta Asp Gly His Ser Pro Ser Pro Thr Ser Glu Ser Glu Pro Asp Leu 745 750 755	2370
gaa tcc cag tat cca ggc tca ggg agt att cct cca aac cag ccc tct Glu Ser Gln Tyr Pro Gly Ser Gly Ser Ile Pro Pro Asn Gln Pro Ser 760 765 770 775	2418
ggt gac cct cag cag ccc agc cct gac agt act gct gcc cag aaa gtt Gly Asp Pro Gln Gln Pro Ser Pro Asp Ser Thr Ala Ala Gln Lys Val 780 785 790	2466
gcc aca agt ccc aag agt gcc ctc aag tct cca tct tcc aag cgt agg Ala Thr Ser Pro Lys Ser Ala Leu Lys Ser Pro Ser Ser Lys Arg Arg 795 800 805	2514
aca tct cag aac tta aaa ctg aga gtt acc ttt gag gag cct gtg gtg Thr Ser Gln Asn Leu Lys Leu Arg Val Thr Phe Glu Glu Pro Val Val 810 815 820	2562
cag atg gag cag cct agc ctt gaa ctg aat gga gaa aaa gac aaa gat Gln Met Glu Gln Pro Ser Leu Glu Leu Asn Gly Glu Lys Asp Lys Asp 825 830 835	2610
aag ggc agg act ctc cag cgg acc tcc aca agt aac gaa tcg ggg gat Lys Gly Arg Thr Leu Gln Arg Thr Ser Thr Ser Asn Glu Ser Gly Asp 840 845 850 855	2658
caa ctg aaa agg cct ttt gga gcc ttt cga tct atc atg gag aca cta Gln Leu Lys Arg Pro Phe Gly Ala Phe Arg Ser Ile Met Glu Thr Leu 860 865 870	2706
agt ggc aac caa aac aat aat aat aac tac cag gca gcc aac cag ctg Ser Gly Asn Gln Asn Asn Asn Asn Asn Tyr Gln Ala Ala Asn Gln Leu 875 880 885	2754
aaa acc tct aca ttg ccc ttg acc tca ctt ggg agg aag aca gat gcc Lys Thr Ser Thr Leu Pro Leu Thr Ser Leu Gly Arg Lys Thr Asp Ala 890 895 900	2802
aag gga aac cct gcc agc tcc gct agc aaa gga aag aat aag gca gca Lys Gly Asn Pro Ala Ser Ser Ala Ser Lys Gly Lys Asn Lys Ala Ala 905 910 915	2850
taatgacatc aatagaaaaa tgaagaaatc ctacagcata aagcacattg ctgagccaga	2910
gtcaaaagaa ctcttcttgt aaatcacttt ttaaattttc tctcactgat gccctttgga	2970
aattattgga aatttctgga ctatcctctt tggaagaga accatgaaaa caatgectca	3030
ccagcagaag aacagaatat caggatgcct taaatttata gtagtagact gtaaaagatt	3090
cattttgggg tgatatctgt atatataact tgttttttta aaagatgccg tttaaaagca	3150
tgattgggaa aatgtacgtt ttttaagagt agattgattc accctacca caggacattc	3210

accaagccac tgataccatt ttatatttca tcaattgcat gagtatttgc taatgttgat 3270  
 tgaacctccc tttocccata atgtgggcag atttggctca gctccttcat gagatcaggt 3330  
 cagtgggtatt gtttctgtca agagtgtttt ttctgtcatt tctacttttt gtataaagga 3390  
 aataaaacaa tgtaacagc caaaaaaaaa aaaaaaaaaa .aa 3432

<210> 2

<211> 919

<212> PRT

<213> NM\_005460 SNCAIP, synphilin

<400> 2

Met Glu Ala Pro Glu Tyr Leu Asp Leu Asp Glu Ile Asp Phe Ser Asp  
1 5 10 15

Asp Ile Ser Tyr Ser Val Thr Ser Leu Lys Thr Ile Pro Glu Leu Cys  
20 25 30

Arg Arg Cys Asp Thr Gln Asn Glu Asp Arg Ser Ala Ser Ser Ser Ser  
35 40 45

Trp Asn Cys Gly Ile Ser Thr Leu Ile Thr Asn Thr Gln Lys Pro Thr  
50 55 60

Gly Ile Ala Asp Val Tyr Ser Lys Phe Arg Pro Val Lys Arg Val Ser  
65 70 75 80

Pro Leu Lys His Gln Pro Glu Thr Leu Glu Asn Asn Glu Ser Asp Asp  
85 90 95

Gln Lys Asn Gln Lys Val Val Glu Tyr Gln Lys Gly Gly Glu Ser Asp  
100 105 110

Leu Gly Pro Gln Pro Gln Glu Leu Gly Pro Gly Asp Gly Val Gly Gly  
115 120 125

Pro Pro Gly Lys Ser Ser Glu Pro Ser Thr Ser Leu Gly Glu Leu Glu  
130 135 140

His Tyr Asp Leu Asp Met Asp Glu Ile Leu Asp Val Pro Tyr Ile Lys  
145 150 155 160

Ser Ser Gln Gln Leu Ala Ser Phe Thr Lys Val Thr Ser Glu Lys Arg  
165 170 175

Ile Leu Gly Leu Cys Thr Thr Ile Asn Gly Leu Ser Gly Lys Ala Cys  
180 185 190

Ser Thr Gly Ser Ser Glu Ser Ser Ser Ser Asn Met Ala Pro Phe Cys  
195 200 205

Val Leu Ser Pro Val Lys Ser Pro His Leu Arg Lys Ala Ser Ala Val  
210 215 220

Ile His Asp Gln His Lys Leu Ser Thr Glu Glu Thr Glu Ile Ser Pro  
225 230 235 240

Pro Leu Val Lys Cys Gly Ser Ala Tyr Glu Pro Glu Asn Gln Ser Lys  
245 250 255

Asp Phe Leu Asn Lys Thr Phe Ser Asp Pro His Gly Arg Lys Val Glu  
260 265 270

Lys Thr Thr Pro Asp Cys Gln Leu Arg Ala Phe His Leu Gln Ser Ser  
275 280 285

Ala Ala Glu Ser Lys Pro Glu Glu Gln Val Ser Gly Leu Asn Arg Thr  
290 295 300

Ser Ser Gln Gly Pro Glu Glu Arg Ser Glu Tyr Leu Lys Lys Val Lys  
305 310 315 320

Ser Ile Leu Asn Ile Val Lys Glu Gly Gln Ile Ser Leu Leu Pro His  
325 330 335

Leu Ala Ala Asp Asn Leu Asp Lys Ile His Asp Glu Asn Gly Asn Asn  
340 345 350

Leu Leu His Ile Ala Ala Ser Gln Gly His Ala Glu Cys Leu Gln His  
355 360 365

Leu Thr Ser Leu Met Gly Glu Asp Cys Leu Asn Glu Arg Asn Thr Glu  
370 375 380

Lys Leu Thr Pro Ala Gly Leu Ala Ile Lys Asn Gly Gln Leu Glu Cys  
385 390 395 400

Val Arg Trp Met Val Ser Glu Thr Glu Ala Ile Ala Glu Leu Ser Cys  
405 410 415

Ser Lys Asp Phe Pro Ser Leu Ile His Tyr Ala Gly Cys Tyr Gly Gln  
420 425 430

Glu Lys Ile Leu Leu Trp Leu Leu Gln Phe Met Gln Glu Gln Gly Ile  
435 440 445

Ser Leu Asp Glu Val Asp Gln Asp Gly Asn Ser Ala Val His Val Ala  
450 455 460

Ser Gln His Gly Tyr Leu Gly Cys Ile Gln Thr Leu Val Glu Tyr Gly  
 465 470 475 480

Ala Asn Val Thr Met Gln Asn His Ala Gly Glu Lys Pro Ser Gln Ser  
 485 490 495

Ala Glu Arg Gln Gly His Thr Leu Cys Ser Arg Tyr Leu Val Val Val  
 500 505 510

Glu Thr Cys Met Ser Leu Ala Ser Gln Val Val Lys Leu Thr Lys Gln  
 515 520 525

Leu Lys Glu Gln Thr Val Glu Arg Val Thr Leu Gln Asn Gln Leu Gln  
 530 535 540

Gln Phe Leu Glu Ala Gln Lys Ser Glu Gly Lys Ser Leu Pro Ser Ser  
 545 550 555 560

Pro Ser Ser Pro Ser Ser Pro Ala Ser Arg Lys Ser Gln Trp Lys Ser  
 565 570 575

Pro Asp Ala Asp Asp Asp Ser Val Ala Lys Ser Lys Pro Gly Val Gln  
 580 585 590

Glu Gly Ile Gln Val Leu Gly Ser Leu Ser Ala Ser Ser Arg Ala Arg  
 595 600 605

Pro Lys Ala Lys Asp Glu Asp Ser Asp Lys Ile Leu Arg Gln Leu Leu  
 610 615 620

Gly Lys Glu Ile Ser Glu Asn Val Cys Thr Gln Glu Lys Leu Ser Leu  
 625 630 635 640

Glu Phe Gln Asp Ala Gln Ala Ser Ser Arg Asn Ser Lys Lys Ile Pro  
 645 650 655

Leu Glu Lys Arg Glu Leu Lys Leu Ala Arg Leu Arg Gln Leu Met Gln  
 660 665 670

Arg Ser Leu Ser Glu Ser Asp Thr Asp Ser Asn Asn Ser Glu Asp Pro  
 675 680 685

Lys Thr Thr Pro Val Arg Lys Ala Asp Arg Pro Arg Pro Gln Pro Ile  
 690 695 700

Val Glu Ser Val Glu Ser Met Asp Ser Ala Glu Ser Leu His Leu Met  
 705 710 715 720

Ile Lys Lys His Thr Leu Ala Ser Gly Gly Arg Arg Phe Pro Phe Ser  
 725 730 735

Ile Lys Ala Ser Lys Ser Leu Asp Gly His Ser Pro Ser Pro Thr Ser  
740 745 750

Glu Ser Ser Glu Pro Asp Leu Glu Ser Gln Tyr Pro Gly Ser Gly Ser  
755 760 765

Ile Pro Pro Asn Gln Pro Ser Gly Asp Pro Gln Gln Pro Ser Pro Asp  
770 775 780

Ser Thr Ala Ala Gln Lys Val Ala Thr Ser Pro Lys Ser Ala Leu Lys  
785 790 795 800

Ser Pro Ser Ser Lys Arg Arg Thr Ser Gln Asn Leu Lys Leu Arg Val  
805 810 815

Thr Phe Glu Glu Pro Val Val Gln Met Glu Gln Pro Ser Leu Glu Leu  
820 825 830

Asn Gly Glu Lys Asp Lys Asp Lys Gly Arg Thr Leu Gln Arg Thr Ser  
835 840 845

Thr Ser Asn Glu Ser Gly Asp Gln Leu Lys Arg Pro Phe Gly Ala Phe  
850 855 860

Arg Ser Ile Met Glu Thr Leu Ser Gly Asn Gln Asn Asn Asn Asn  
865 870 875 880

Tyr Gln Ala Ala Asn Gln Leu Lys Thr Ser Thr Leu Pro Leu Thr Ser  
885 890 895

Leu Gly Arg Lys Thr Asp Ala Lys Gly Asn Pro Ala Ser Ser Ala Ser  
900 905 910

Lys Gly Lys Asn Lys Ala Ala  
915

<210> 3

<211> 4181

<212> DNA

<213> NM\_002387 MCC, mutated in colorectal cancers

<220>

<221> CDS

<222> (221) .. (2707)

<223>

<400> 3	
cctcctgcag caatggctcg tccgtgaaac gcgagccacg gctgctcttt ttaagagtgc	60
ctgcatcctc cgtttgcgct tcgcaactgt cctgggtgaa aatggctgtc tagactaaaa	120
tgtggcagaa gggaccaagc agtgatatt gagcctgtga agtccaactc ttaagctccg	180
agacctgggg gactgagagc ccagctctga aaagtgcac atg aat tcc gga gtt	235
Met Asn Ser Gly Val	
1 5	
gcc atg aaa tat gga aac gac tcc tcg gcc gag ctg agt gag ctc cat	283
Ala Met Lys Tyr Gly Asn Asp Ser Ser Ala Glu Leu Ser Glu Leu His	
10 15 20	
tca gca gcc ctg gca tca cta aag gga gat ata gtg gaa ctt aat aaa	331
Ser Ala Ala Leu Ala Ser Leu Lys Gly Asp Ile Val Glu Leu Asn Lys	
25 30 35	
cgt ctc cag caa aca gag agg gaa cgg gac ctt ctg gaa aag aaa ttg	379
Arg Leu Gln Gln Thr Glu Arg Glu Arg Asp Leu Leu Glu Lys Lys Leu	
40 45 50	
gcc aag gca cag tgc gag cag tcc cac ctc atg aga gag cat gag gat	427
Ala Lys Ala Gln Cys Glu Gln Ser His Leu Met Arg Glu His Glu Asp	
55 60 65	
gtc cag gag cga acg acg ctt cgc tat gag gaa cgc atc aca gag ctc	475
Val Gln Glu Arg Thr Thr Leu Arg Tyr Glu Glu Arg Ile Thr Glu Leu	
70 75 80 85	
cac agc gtc att gcg gag ctc aac aag aag ata gac cgt ctg caa ggc	523
His Ser Val Ile Ala Glu Leu Asn Lys Lys Ile Asp Arg Leu Gln Gly	
90 95 100	
acc acc atc agg gag gaa gat gag tac tca gaa ctg cga tca gaa ctc	571
Thr Thr Ile Arg Glu Glu Asp Glu Tyr Ser Glu Leu Arg Ser Glu Leu	
105 110 115	
agc cag agc caa cac gag gtc aac gag gac tct cga agc atg gac caa	619
Ser Gln Ser Gln His Glu Val Asn Glu Asp Ser Arg Ser Met Asp Gln	
120 125 130	
gac cag acc tct gtc tct atc ccc gaa aac cag tct acc atg gtt act	667
Asp Gln Thr Ser Val Ser Ile Pro Glu Asn Gln Ser Thr Met Val Thr	
135 140 145	
gct gac atg gac aac tgc agt gac ctg aac tca gaa ctg cag agg gtg	715
Ala Asp Met Asp Asn Cys Ser Asp Leu Asn Ser Glu Leu Gln Arg Val	
150 155 160 165	
ctg aca ggg ctg gag aat gtt gtc tgc ggc agg aag aag agc agc tgc	763
Leu Thr Gly Leu Glu Asn Val Val Cys Gly Arg Lys Lys Ser Ser Cys	
170 175 180	
agc ctc tcc gtg gcc gag gtg gac agg cac att gag cag ctc acc aca	811
Ser Leu Ser Val Ala Glu Val Asp Arg His Ile Glu Gln Leu Thr Thr	
185 190 195	
gcc agc gag cac tgt gac ctg gct att aag aca gtc gag gag att gag	859
Ala Ser Glu His Cys Asp Leu Ala Ile Lys Thr Val Glu Glu Ile Glu	
200 205 210	
ggg gtg ctt ggc cgg gac ctg tat ccc aac ctg gct gaa gag agg tct	907
Gly Val Leu Gly Arg Asp Leu Tyr Pro Asn Leu Ala Glu Glu Arg Ser	
215 220 225	

cgg tgg gag aag gag ctg gct ggg ctg agg gaa gag aat gag agc ctg Arg Trp Glu Lys Glu Leu Ala Gly Leu Arg Glu Glu Asn Glu Ser Leu 230 235 240 245	955
act gcc atg ctg tgc agc aaa gag gaa gaa ctg aac cgg act aag gcc Thr Ala Met Leu Cys Ser Lys Glu Glu Glu Leu Asn Arg Thr Lys Ala 250 255 260	1003
acc atg aat gcc atc cgg gaa gag cgg gac cgg ctc cgg agg cgg gtc Thr Met Asn Ala Ile Arg Glu Glu Arg Asp Arg Leu Arg Arg Arg Val 265 270 275	1051
aga gag ctt caa act cga cta cag agc gtg cag gcc aca ggt ccc tcc Arg Glu Leu Gln Thr Arg Leu Gln Ser Val Gln Ala Thr Gly Pro Ser 280 285 290	1099
agc cct ggc cgc ctc act tcc acc aac cgc cgg att aac ccc agc act Ser Pro Gly Arg Leu Thr Ser Thr Asn Arg Pro Ile Asn Pro Ser Thr 295 300 305	1147
ggg gag ctg agc aca agc agc agc agc aat gac att ccc atc gcc aag Gly Glu Leu Ser Thr Ser Ser Ser Asn Asp Ile Pro Ile Ala Lys 310 315 320 325	1195
att gct gag agg gtg aag cta tca aag aca agg tcc gaa tcg tca tca Ile Ala Glu Arg Val Lys Leu Ser Lys Thr Arg Ser Glu Ser Ser Ser 330 335 340	1243
tct gat cgg cca gtc ctg ggc tca gaa atc agt agc ata ggg gta tcc Ser Asp Arg Pro Val Leu Gly Ser Glu Ile Ser Ser Ile Gly Val Ser 345 350 355	1291
agc agt gtg gct gaa cac ctg gcc cac tca ctt cag gac tgc tcc aat Ser Ser Val Ala Glu His Leu Ala His Ser Leu Gln Asp Cys Ser Asn 360 365 370	1339
atc caa gag att ttc caa aca ctc tac tca cac gga tct gcc atc tca Ile Gln Glu Ile Phe Gln Thr Leu Tyr Ser His Gly Ser Ala Ile Ser 375 380 385	1387
gaa agc aag att aga gag ttt gag gtg gaa aca gaa cgg ctg aat agc Glu Ser Lys Ile Arg Glu Phe Glu Val Glu Thr Glu Arg Leu Asn Ser 390 395 400 405	1435
cgg att gag cac ctc aaa tcc caa aat gac ctc ctg acc ata acc ttg Arg Ile Glu His Leu Lys Ser Gln Asn Asp Leu Leu Thr Ile Thr Leu 410 415 420	1483
gag gaa tgt aaa agc aat gct gag agg atg agc atg ctg gtg gga aaa Glu Glu Cys Lys Ser Asn Ala Glu Arg Met Ser Met Leu Val Gly Lys 425 430 435	1531
tac gaa tcc aat gcc aca gcg ctg agg ctg gcc ttg cag tac agc gag Tyr Glu Ser Asn Ala Thr Ala Leu Arg Leu Ala Leu Gln Tyr Ser Glu 440 445 450	1579
cag tgc atc gaa gcc tac gaa ctc ctc ctg gcg ctg gca gag agt gag Gln Cys Ile Glu Ala Tyr Glu Leu Leu Leu Ala Leu Ala Glu Ser Glu 455 460 465	1627
cag agc ctc atc ctg ggg cag ttc cga gcg gcg ggc gtg ggg tcc tcc Gln Ser Leu Ile Leu Gly Gln Phe Arg Ala Ala Gly Val Gly Ser Ser 470 475 480 485	1675
cct gga gac cag tcg ggg gat gaa aac atc act cag atg ctc aag cga Pro Gly Asp Gln Ser Gly Asp Glu Asn Ile Thr Gln Met Leu Lys Arg 490 495 500	1723



gct cat gac tgc cgg aag aca gct gag aac gct gcc aag gcc ctg ctc Ala His Asp Cys Arg Lys Thr Ala Glu Asn Ala Ala Lys Ala Leu Leu 505 510 515	1771
atg aag ctg gac ggc agc tgt ggg gga gcc ttt gcc gtg gcc ggc tgc Met Lys Leu Asp Gly Ser Cys Gly Gly Ala Phe Ala Val Ala Gly Cys 520 525 530	1819
agc gtg cag ccc tgg gag agc ctt tcc tcc aac agc cac acc agc aca Ser Val Gln Pro Trp Glu Ser Leu Ser Ser Asn Ser His Thr Ser Thr 535 540 545	1867
acc agc tcc aca gcc agt agt tgc gac acc gag ttc act aaa gaa gac Thr Ser Ser Thr Ala Ser Ser Cys Asp Thr Glu Phe Thr Lys Glu Asp 550 555 560 565	1915
gag cag agg ctg aag gat tat atc cag cag ctc aag aat gac agg gct Glu Gln Arg Leu Lys Asp Tyr Ile Gln Gln Leu Lys Asn Asp Arg Ala 570 575 580	1963
gcg gtc aag ctg acc atg ctg gag ctg gaa agc atc cac atc gat cct Ala Val Lys Leu Thr Met Leu Glu Leu Glu Ser Ile His Ile Asp Pro 585 590 595	2011
ctc agc tat gac gtc aag cct cgg gga gac agc cag agg ctg gat ctg Leu Ser Tyr Asp Val Lys Pro Arg Gly Asp Ser Gln Arg Leu Asp Leu 600 605 610	2059
gaa aac gca gtg ctt atg cag gag ctc atg gcc atg aag gag gag atg Glu Asn Ala Val Leu Met Gln Glu Leu Met Ala Met Lys Glu Glu Met 615 620 625	2107
gcc gag ttg aag gcc cag ctc tac cta ctg gag aaa gag aag aag gcc Ala Glu Leu Lys Ala Gln Leu Tyr Leu Leu Glu Lys Glu Lys Lys Ala 630 635 640 645	2155
ctg gag ctg aag ctg agc acg cgg gag gcc cag gag cag gcc tac ctg Leu Glu Leu Lys Leu Ser Thr Arg Glu Ala Gln Glu Gln Ala Tyr Leu 650 655 660	2203
gtg cac att gag cac ctg aag tcc gag gtg gag gag cag aag gag cag Val His Ile Glu His Leu Lys Ser Glu Val Glu Glu Gln Lys Glu Gln 665 670 675	2251
cgg atg cga tcc ctc agc tcc acc agc agc ggc agc aaa gat aaa cct Arg Met Arg Ser Leu Ser Ser Thr Ser Ser Gly Ser Lys Asp Lys Pro 680 685 690	2299
ggc aag gag tgt gct gat gct gcc tcc cca gct ctg tcc cta gct gaa Gly Lys Glu Cys Ala Asp Ala Ala Ser Pro Ala Leu Ser Leu Ala Glu 695 700 705	2347
ctc agg aca acg tgc agc gag aat gag ctg gct gcg gag ttc acc aac Leu Arg Thr Thr Cys Ser Glu Asn Glu Leu Ala Ala Glu Phe Thr Asn 710 715 720 725	2395
gcc att cgt cga gaa aag aag ttg aag gcc aga gtt caa gag ctg gtg Ala Ile Arg Arg Glu Lys Lys Leu Lys Ala Arg Val Gln Glu Leu Val 730 735 740	2443
agt gcc ttg gag aga ctc acc aag agc agt gaa atc cga cat cag caa Ser Ala Leu Glu Arg Leu Thr Lys Ser Ser Glu Ile Arg His Gln Gln 745 750 755	2491
tct gca gag ttc gtg aat gat cta aag cgg gcc aac agc aac ctg gtg Ser Ala Glu Phe Val Asn Asp Leu Lys Arg Ala Asn Ser Asn Leu Val 760 765 770	2539

gct gcc tat gag aaa gca aag aaa aag cat caa aac aaa ctg aag aag 2587  
 Ala Ala Tyr Glu Lys Ala Lys Lys Lys His Gln Asn Lys Leu Lys Lys  
 775 780 785  
 tta gag tcg cag atg atg gcc atg gtg gag aga cat gag acc caa gtg 2635  
 Leu Glu Ser Gln Met Met Ala Met Val Glu Arg His Glu Thr Gln Val  
 790 795 800 805  
 agg atg ctc aag caa aga ata gct ctg cta gag gag gag aac tcc agg 2683  
 Arg Met Leu Lys Gln Arg Ile Ala Leu Leu Glu Glu Glu Asn Ser Arg  
 810 815 820  
 cca cac acc aat gaa act tcg ctt taatcagcac tcacgcaccg gagttctgcc 2737  
 Pro His Thr Asn Glu Thr Ser Leu  
 825  
 catgggaagt aaactgcagc aggccactgg ggacagaagg gcccattgtac ttgttgggag 2797  
 gagggaggaaa gggaaggctg gcaggtaggt cggcacttgg acaatggagt gcccgaactc 2857  
 aacccttggg gtgactggcc atggtgacat tgtggactgt atccagaggt gcccgtcttt 2917  
 cctccttggg cccacaacag cgtgtaaaca catgttctgt gctgtctcag cagagcctcg 2977  
 tttctgcttt cagcactcac tctccccctc ctcttctggt ctggcggtg tgcattcagt 3037  
 ggatcccaga catttgtttc tgtaagattt tccattgtat cctctttttg gtagatgctg 3097  
 ggctcatctt ctagaatctc gtttctctc tttcctctg ctccatggga aaacagacct 3157  
 gtgtgtgcct ccagcattta aaaggactgc tgatttgttt actacagcaa ggctttgggt 3217  
 tccaagtccc ggggtctcaac tttaagatag aggcggccat aagaggatgat ctctgggagt 3277  
 tataggatcat gggaagagcg tagacaggtg ttacttacag toccagatac actaaagtta 3337  
 caaacagacc accaccagga ctgtgcctga acaattttgt attgagagaa taaaaacttc 3397  
 ctccaatctt cattttggag gcagggtgga gaaggagcg ctctcttgat tctgggattt 3457  
 ctccctctca gtggagcctt attaatatcc aagacttaga gctgggaatc tttttgatac 3517  
 ctgtagtga actaaaattc tgtcaggggt ttcttcaaga gctgagaaac attattagca 3577  
 ctccccccc cagggcacta cataattgct gttctgctga atcaaatctc ttccacatgg 3637  
 gtgcatttgt agctctggac ctgtctctac ctaaggacaa gacactgagg agatactgaa 3697  
 cattttgcaa aacttatcac gctacttaa gactgctgtg taacccccag ttcaagactt 3757  
 agctcctgtt gtcattgacg ggacagagtg agggaatggt agttaaggct tcttttttgc 3817  
 cccagatac atggtgatgg ttagcatatg gtgcttaaaa ggttaaattt caagcaaaat 3877  
 gcttacaggg ctaggcagta ccaaagtaac tgaattattt caggaagggtc ttcaatctta 3937  
 aaacaaattc attattcttt ttcatgttta cctcttctct ctcatgttca cactgataca 3997  
 ctigaaggac catttactgt tttttctgt agcaccagag aatccatcca aagttcccta 4057  
 tgaaaaatgt gttccattgc catagctgac tacaaattaa agttgaggag gtttctgcat 4117  
 agagtcttta tgtccataag ctacgggtag gtctattttc agagcatgat acaaatcca 4177  
 cagg 4181

&lt;210&gt; 4

&lt;211&gt; 829

&lt;212&gt; PRT

&lt;213&gt; NM\_002387 MCC, mutated in colorectal cancers

&lt;400&gt; 4

Met Asn Ser Gly Val Ala Met Lys Tyr Gly Asn Asp Ser Ser Ala Glu  
 1 5 10 15

Leu Ser Glu Leu His Ser Ala Ala Leu Ala Ser Leu Lys Gly Asp Ile  
 20 25 30

Val Glu Leu Asn Lys Arg Leu Gln Gln Thr Glu Arg Glu Arg Asp Leu  
 35 40 45

Leu Glu Lys Lys Leu Ala Lys Ala Gln Cys Glu Gln Ser His Leu Met  
 50 55 60

Arg Glu His Glu Asp Val Gln Glu Arg Thr Thr Leu Arg Tyr Glu Glu  
 65 70 75 80

Arg Ile Thr Glu Leu His Ser Val Ile Ala Glu Leu Asn Lys Lys Ile  
 85 90 95

Asp Arg Leu Gln Gly Thr Thr Ile Arg Glu Glu Asp Glu Tyr Ser Glu  
 100 105 110

Leu Arg Ser Glu Leu Ser Gln Ser Gln His Glu Val Asn Glu Asp Ser  
 115 120 125

Arg Ser Met Asp Gln Asp Gln Thr Ser Val Ser Ile Pro Glu Asn Gln  
 130 135 140

Ser Thr Met Val Thr Ala Asp Met Asp Asn Cys Ser Asp Leu Asn Ser  
 145 150 155 160

Glu Leu Gln Arg Val Leu Thr Gly Leu Glu Asn Val Val Cys Gly Arg  
 165 170 175

Lys Lys Ser Ser Cys Ser Leu Ser Val Ala Glu Val Asp Arg His Ile  
 180 185 190

Glu Gln Leu Thr Thr Ala Ser Glu His Cys Asp Leu Ala Ile Lys Thr  
 195 200 205

Val Glu Glu Ile Glu Gly Val Leu Gly Arg Asp Leu Tyr Pro Asn Leu  
 210 215 220

Ala Glu Glu Arg Ser Arg Trp Glu Lys Glu Leu Ala Gly Leu Arg Glu  
225 230 235 240

Glu Asn Glu Ser Leu Thr Ala Met Leu Cys Ser Lys Glu Glu Glu Leu  
245 250 255

Asn Arg Thr Lys Ala Thr Met Asn Ala Ile Arg Glu Glu Arg Asp Arg  
260 265 270

Leu Arg Arg Arg Val Arg Glu Leu Gln Thr Arg Leu Gln Ser Val Gln  
275 280 285

Ala Thr Gly Pro Ser Ser Pro Gly Arg Leu Thr Ser Thr Asn Arg Pro  
290 295 300

Ile Asn Pro Ser Thr Gly Glu Leu Ser Thr Ser Ser Ser Ser Asn Asp  
305 310 315 320

Ile Pro Ile Ala Lys Ile Ala Glu Arg Val Lys Leu Ser Lys Thr Arg  
325 330 335

Ser Glu Ser Ser Ser Ser Asp Arg Pro Val Leu Gly Ser Glu Ile Ser  
340 345 350

Ser Ile Gly Val Ser Ser Ser Val Ala Glu His Leu Ala His Ser Leu  
355 360 365

Gln Asp Cys Ser Asn Ile Gln Glu Ile Phe Gln Thr Leu Tyr Ser His  
370 375 380

Gly Ser Ala Ile Ser Glu Ser Lys Ile Arg Glu Phe Glu Val Glu Thr  
385 390 395 400

Glu Arg Leu Asn Ser Arg Ile Glu His Leu Lys Ser Gln Asn Asp Leu  
405 410 415

Leu Thr Ile Thr Leu Glu Glu Cys Lys Ser Asn Ala Glu Arg Met Ser  
420 425 430

Met Leu Val Gly Lys Tyr Glu Ser Asn Ala Thr Ala Leu Arg Leu Ala  
435 440 445

Leu Gln Tyr Ser Glu Gln Cys Ile Glu Ala Tyr Glu Leu Leu Leu Ala  
450 455 460

Leu Ala Glu Ser Glu Gln Ser Leu Ile Leu Gly Gln Phe Arg Ala Ala  
465 470 475 480

Gly Val Gly Ser Ser Pro Gly Asp Gln Ser Gly Asp Glu Asn Ile Thr  
485 490 495

Gln Met Leu Lys Arg Ala His Asp Cys Arg Lys Thr Ala Glu Asn Ala  
500 505 510

Ala Lys Ala Leu Leu Met Lys Leu Asp Gly Ser Cys Gly Gly Ala Phe  
515 520 525

Ala Val Ala Gly Cys Ser Val Gln Pro Trp Glu Ser Leu Ser Ser Asn  
530 535 540

Ser His Thr Ser Thr Thr Ser Ser Thr Ala Ser Ser Cys Asp Thr Glu  
545 550 555 560

Phe Thr Lys Glu Asp Glu Gln Arg Leu Lys Asp Tyr Ile Gln Gln Leu  
565 570 575

Lys Asn Asp Arg Ala Ala Val Lys Leu Thr Met Leu Glu Leu Glu Ser  
580 585 590

Ile His Ile Asp Pro Leu Ser Tyr Asp Val Lys Pro Arg Gly Asp Ser  
595 600 605

Gln Arg Leu Asp Leu Glu Asn Ala Val Leu Met Gln Glu Leu Met Ala  
610 615 620

Met Lys Glu Glu Met Ala Glu Leu Lys Ala Gln Leu Tyr Leu Leu Glu  
625 630 635 640

Lys Glu Lys Lys Ala Leu Glu Leu Lys Leu Ser Thr Arg Glu Ala Gln  
645 650 655

Glu Gln Ala Tyr Leu Val His Ile Glu His Leu Lys Ser Glu Val Glu  
660 665 670

Glu Gln Lys Glu Gln Arg Met Arg Ser Leu Ser Ser Thr Ser Ser Gly  
675 680 685

Ser Lys Asp Lys Pro Gly Lys Glu Cys Ala Asp Ala Ala Ser Pro Ala  
690 695 700

Leu Ser Leu Ala Glu Leu Arg Thr Thr Cys Ser Glu Asn Glu Leu Ala  
705 710 715 720

Ala Glu Phe Thr Asn Ala Ile Arg Arg Glu Lys Lys Leu Lys Ala Arg  
725 730 735

Val Gln Glu Leu Val Ser Ala Leu Glu Arg Leu Thr Lys Ser Ser Glu  
740 745 750

Ile Arg His Gln Gln Ser Ala Glu Phe Val Asn Asp Leu Lys Arg Ala  
755 760 765

Asn Ser Asn Leu Val Ala Ala Tyr Glu Lys Ala Lys Lys Lys His Gln  
770 775 780

Asn Lys Leu Lys Lys Leu Glu Ser Gln Met Met Ala Met Val Glu Arg  
785 790 795 800

His Glu Thr Gln Val Arg Met Leu Lys Gln Arg Ile Ala Leu Leu Glu  
805 810 815

Glu Glu Asn Ser Arg Pro His Thr Asn Glu Thr Ser Leu  
820 825

<210> 5

<211> 2830

<212> DNA

<213> NM\_022117 SE20-4

<220>

<221> CDS

<222> (130)..(2208)

<223>

<400> 5

aattcggcac gaggagagct gggtgcgtga gtctcctcag ctctgcttac cgggtgcgact 60

agcggcagcg acgcggtctaa aagcgaaggg gcgagtgcca gtcccctgag ctgtacgaac 120

gcggtcgcc atg gac cgc cca gat gag ggg cct ccg gcc aag acc cgc cgc 171  
Met Asp Arg Pro Asp Glu Gly Pro Pro Ala Lys Thr Arg Arg  
1 5 10

ctg agc agc tcc gag tct cca cag cgc gac ccg ccc ccg ccg ccg 219  
Leu Ser Ser Ser Glu Ser Pro Gln Arg Asp Pro Pro Pro Pro Pro Pro  
15 20 25 30

ccg ccg ccg ctc ctc cga ctg ccg ctg cct cca ccc cag cag cgc ccg 267  
Pro Pro Pro Leu Leu Arg Leu Pro Leu Pro Pro Pro Gln Gln Arg Pro  
35 40 45

agg ctc cag gag gaa acg gag gcg gca cag gtg ctg gcc gat atg agg 315  
Arg Leu Gln Glu Glu Thr Glu Ala Ala Gln Val Leu Ala Asp Met Arg  
50 55 60

ggg gtg gga ctg ggc ccc gcg ctg ccc ccg ccg cct ccc tat gtc att 363  
Gly Val Gly Leu Gly Pro Ala Leu Pro Pro Pro Pro Tyr Val Ile  
65 70 75

ctc gag gag ggg ggg atc cgc gca tac ttc acg ctc ggt gct gag tgt 411  
Leu Glu Glu Gly Gly Ile Arg Ala Tyr Phe Thr Leu Gly Ala Glu Cys  
80 85 90

ccc gcc tgg gat tct acc atc gag tcg ggg tat ggg gag gcg ccc ccg 459  
Pro Gly Trp Asp Ser Thr Ile Glu Ser Gly Tyr Gly Glu Ala Pro Pro

95	100	105	110	
ccc acg gag agc ctg gaa gca ctc ccc act cct gag gcc tcg ggg ggg				507
Pro Thr Glu Ser Leu Glu Ala Leu Pro Thr Glu Ala Ser Gly Gly	115	120	125	
agc ctg gaa atc gat ttt cag gtt gta cag tcg agc agt ttt ggt gga				555
Ser Leu Glu Ile Asp Phe Gln Val Gln Ser Ser Ser Phe Gly Gly	130	135	140	
gag ggg gcc cta gaa acc tgt agc gca gtg ggg tgg gcg ccc cag agg				603
Glu Gly Ala Leu Glu Thr Cys Ser Ala Val Gly Trp Ala Pro Gln Arg	145	150	155	
tta gtt gac ccg aag agc aag gaa gag gcg atc atc ata gtg gag gat				651
Leu Val Asp Pro Lys Ser Lys Glu Glu Ala Ile Ile Ile Val Glu Asp	160	165	170	
gag gat gag gat gag cgg gag agt atg agg agc agc agg agg cgg cgg				699
Glu Asp Glu Asp Glu Arg Glu Ser Met Arg Ser Ser Arg Arg Arg Arg	175	180	185	190
cgg cgg cgg agg agg aag cag agg aag gtg aag agg gaa agc aga gag				747
Arg Arg Arg Arg Arg Lys Gln Arg Lys Val Lys Arg Glu Ser Arg Glu	195	200	205	
aga aat gcc gag agg atg gag agc atc ctg cag gca ctg gag gat att				795
Arg Asn Ala Glu Arg Met Glu Ser Ile Leu Gln Ala Leu Glu Asp Ile	210	215	220	
cag ctg gat ctg gag gca gtg aac atc aag gca ggc aaa gcc ttc ctg				843
Gln Leu Asp Leu Glu Ala Val Asn Ile Lys Ala Gly Lys Ala Phe Leu	225	230	235	
cgt ctc aag cgc aag ttc atc cag atg cga aga ccc ttc ctg gag cgc				891
Arg Leu Lys Arg Lys Phe Ile Gln Met Arg Arg Pro Phe Leu Glu Arg	240	245	250	
aga gac ctc atc atc cag cat atc cca ggc ttc tgg gtc aaa gca ttc				939
Arg Asp Leu Ile Ile Gln His Ile Pro Gly Phe Trp Val Lys Ala Phe	255	260	265	270
ctc aac cac ccc aga att tca att ttg atc aac cga cgt gat gaa gac				987
Leu Asn His Pro Arg Ile Ser Ile Leu Ile Asn Arg Arg Asp Glu Asp	275	280	285	
att ttc cgc tac ttg acc aat ctg cag gta cag gat ctc aga cat atc				1035
Ile Phe Arg Tyr Leu Thr Asn Leu Gln Val Gln Asp Leu Arg His Ile	290	295	300	
tcc atg ggc tac aaa atg aag ctg tac ttc cag act aac ccc tac ttc				1083
Ser Met Gly Tyr Lys Met Lys Leu Tyr Phe Gln Thr Asn Pro Tyr Phe	305	310	315	
aca aac atg gtg att gtc aag gag ttc cag cgc aac cgc tca ggc cgg				1131
Thr Asn Met Val Ile Val Lys Glu Phe Gln Arg Asn Arg Ser Gly Arg	320	325	330	
ctg gtg tct cac tca acc cca atc cgc tgg cac cgg ggc cag gaa ccc				1179
Leu Val Ser His Ser Thr Pro Ile Arg Trp His Arg Gly Gln Glu Pro	335	340	345	350
cag gcc cgt cgt cac ggg aac cag gat gcg agc cac agc ttt ttc agc				1227
Gln Ala Arg Arg His Gly Asn Gln Asp Ala Ser His Ser Phe Phe Ser	355	360	365	
tgg ttc tca aac cat agc ctc cca gag gct gac agg att gct gag att				1275
Trp Phe Ser Asn His Ser Leu Pro Glu Ala Asp Arg Ile Ala Glu Ile				

370	375	380	
atc aag aat gat ctg tgg gtt aac cct cta cgc tac tac ctg aga gaa Ile Lys Asn Asp Leu Trp Val Asn Pro Leu Arg Tyr Tyr Leu Arg Glu 385 390 395			1323
agg ggc tcc agg ata aag aga aag aag caa gaa atg aag aaa cgt aaa Arg Gly Ser Arg Ile Lys Arg Lys Lys Gln Glu Met Lys Lys Arg Lys 400 405 410			1371
acc agg ggc aga tgt gag gtg gtg atc atg gaa gac gcc cct gac tat Thr Arg Gly Arg Cys Glu Val Val Ile Met Glu Asp Ala Pro Asp Tyr 415 420 425 430			1419
tat gca gtg gaa gac att ttc agc gag atc tca gac att gat gag aca Tyr Ala Val Glu Asp Ile Phe Ser Glu Ile Ser Asp Ile Asp Glu Thr 435 440 445			1467
att cat gac atc aag atc tct gac ttc atg gag acc acc gac tac ttc Ile His Asp Ile Lys Ile Ser Asp Phe Met Glu Thr Thr Asp Tyr Phe 450 455 460			1515
gag acc act gac aat gag ata act gac atc aat gag aac atc tgc gac Glu Thr Thr Asp Asn Glu Ile Thr Asp Ile Asn Glu Asn Ile Cys Asp 465 470 475			1563
agc gag aat cct gac cac aat gag gtc ccc aac aac gag acc act gat Ser Glu Asn Pro Asp His Asn Glu Val Pro Asn Asn Glu Thr Thr Asp 480 485 490			1611
aac aac gag agt gct gat gac cac gaa acc act gac aac aat gag agt Asn Asn Glu Ser Ala Asp Asp His Glu Thr Thr Asp Asn Asn Glu Ser 495 500 505 510			1659
gca gat gac aac aac gag aat cct gaa gac aat aac aag aac act gat Ala Asp Asp Asn Asn Glu Asn Pro Glu Asp Asn Asn Lys Asn Thr Asp 515 520 525			1707
gac aac gaa gag aac cct aac aac aac gag aac act tac ggc aac aac Asp Asn Glu Glu Asn Pro Asn Asn Asn Glu Asn Thr Tyr Gly Asn Asn 530 535 540			1755
ttc ttc aaa ggt ggc ttc tgg ggc agc cat ggc aac aac cag gac agc Phe Phe Lys Gly Gly Phe Trp Gly Ser His Gly Asn Asn Gln Asp Ser 545 550 555			1803
agc gac agt gac aat gaa gca gat gag gcc agt gat gat gaa gat aat Ser Asp Ser Asp Asn Glu Ala Asp Glu Ala Ser Asp Asp Glu Asp Asn 560 565 570			1851
gat ggc aac gaa ggt gac aat gag ggc agt gat gat gat ggc aat gaa Asp Gly Asn Glu Gly Asp Asn Glu Gly Ser Asp Asp Asp Gly Asn Glu 575 580 585 590			1899
ggt gac aat gaa ggc agc gat gat gac gac aga gac att gag tac tat Gly Asp Asn Glu Gly Ser Asp Asp Asp Arg Asp Ile Glu Tyr Tyr 595 600 605			1947
gag aaa gtt att gaa gac ttt gac aag gat cag gct gac tac gag gac Glu Lys Val Ile Glu Asp Phe Asp Lys Asp Gln Ala Asp Tyr Glu Asp 610 615 620			1995
gtg ata gag atc atc tca gac gaa tca gtg gaa gaa gag ggc att gag Val Ile Glu Ile Ile Ser Asp Glu Ser Val Glu Glu Glu Gly Ile Glu 625 630 635			2043
gaa ggc atc cag caa gat gag gac atc tat gag gaa gga aac tat gag Glu Gly Ile Gln Gln Asp Glu Asp Ile Tyr Glu Glu Gly Asn Tyr Glu			2091



640 645 650

gag gaa gga agt gaa gat gtc tgg gaa gaa ggg gaa gat tcg gac gac 2139  
 Glu Glu Gly Ser Glu Asp Val Trp Glu Glu Gly Glu Asp Ser Asp Asp  
 655 660 665 670

tct gac cta gag gat gtg ctt cag gtc cca aac ggt tgg gcc aat ccg 2187  
 Ser Asp Leu Glu Asp Val Leu Gln Val Pro Asn Gly Trp Ala Asn Pro  
 675 680 685

ggg aag agg ggg aaa acc gga taagggtttt ccccttttgg ggatcacctc 2238  
 Gly Lys Arg Gly Lys Thr Gly  
 690

tctgtatccc ccaccacta tcccatttgc cctcctcctc agctagggcc acgcggcccc 2298

acattgcact tctggggggt gaccgacttc gtacacgggt ttaaagttaa tttttatggt 2358

ttagtcattg cagagttctt attttggggg gagggaaagg gggctagtcc cttcttttg 2418

gccctccgcc ccgcaggct tctgtgtgct gctaactgta tttattgtga tgccttggtc 2478

agggccctc taccacttc tccagtcag ttgtggcccc agcccctctc cctgtgtgt 2538

gtggagtga caccctgacc ccgaagcgg ggaggccgc tgtggccttc gtcacagccg 2598

cgcagtgc atggaggcgc tgtgccacc ttcctctccc aagttcttcc tccatccctc 2658

tcctcttccc gccgcgcgc tagccgcct cgggtgtctat gcaaggccgc ttcgccattg 2718

cggattctt tgcggtattc ttgtcccggt ccccccagaag gctgcctct ccccggtggac 2778

cctgttaatc ccaataaaat tctgagcaag ttcaaaaaaa aaaaaaaaaa aa 2830

&lt;210&gt; 6

&lt;211&gt; 693

&lt;212&gt; PRT

&lt;213&gt; NM\_022117 SE20-4

&lt;400&gt; 6

Met Asp Arg Pro Asp Glu Gly Pro Pro Ala Lys Thr Arg Arg Leu Ser  
 1 5 10 15

Ser Ser Glu Ser Pro Gln Arg Asp Pro Pro Pro Pro Pro Pro Pro  
 20 25 30

Pro Leu Leu Arg Leu Pro Leu Pro Pro Pro Gln Gln Arg Pro Arg Leu  
 35 40 45

Gln Glu Glu Thr Glu Ala Ala Gln Val Leu Ala Asp Met Arg Gly Val  
 50 55 60

Gly Leu Gly Pro Ala Leu Pro Pro Pro Pro Tyr Val Ile Leu Glu  
 65 70 75 80

Glu Gly Gly Ile Arg Ala Tyr Phe Thr Leu Gly Ala Glu Cys Pro Gly

85 90 95

Trp Asp Ser Thr Ile Glu Ser Gly Tyr Gly Glu Ala Pro Pro Pro Thr  
100 105 110

Glu Ser Leu Glu Ala Leu Pro Thr Pro Glu Ala Ser Gly Gly Ser Leu  
115 120 125

Glu Ile Asp Phe Gln Val Val Gln Ser Ser Ser Phe Gly Gly Glu Gly  
130 135 140

Ala Leu Glu Thr Cys Ser Ala Val Gly Trp Ala Pro Gln Arg Leu Val  
145 150 155 160

Asp Pro Lys Ser Lys Glu Glu Ala Ile Ile Ile Val Glu Asp Glu Asp  
165 170 175

Glu Asp Glu Arg Glu Ser Met Arg Ser Ser Arg Arg Arg Arg Arg Arg  
180 185 190

Arg Arg Arg Lys Gln Arg Lys Val Lys Arg Glu Ser Arg Glu Arg Asn  
195 200 205

Ala Glu Arg Met Glu Ser Ile Leu Gln Ala Leu Glu Asp Ile Gln Leu  
210 215 220

Asp Leu Glu Ala Val Asn Ile Lys Ala Gly Lys Ala Phe Leu Arg Leu  
225 230 235 240

Lys Arg Lys Phe Ile Gln Met Arg Arg Pro Phe Leu Glu Arg Arg Asp  
245 250 255

Leu Ile Ile Gln His Ile Pro Gly Phe Trp Val Lys Ala Phe Leu Asn  
260 265 270

His Pro Arg Ile Ser Ile Leu Ile Asn Arg Arg Asp Glu Asp Ile Phe  
275 280 285

Arg Tyr Leu Thr Asn Leu Gln Val Gln Asp Leu Arg His Ile Ser Met  
290 295 300

Gly Tyr Lys Met Lys Leu Tyr Phe Gln Thr Asn Pro Tyr Phe Thr Asn  
305 310 315 320

Met Val Ile Val Lys Glu Phe Gln Arg Asn Arg Ser Gly Arg Leu Val  
325 330 335

Ser His Ser Thr Pro Ile Arg Trp His Arg Gly Gln Glu Pro Gln Ala  
340 345 350

Arg Arg His Gly Asn Gln Asp Ala Ser His Ser Phe Phe Ser Trp Phe

355                      360                      365  
 Ser Asn His Ser Leu Pro Glu Ala Asp Arg Ile Ala Glu Ile Ile Lys  
 370                      375                      380  
 Asn Asp Leu Trp Val Asn Pro Leu Arg Tyr Tyr Leu Arg Glu Arg Gly  
 385                      390                      395                      400  
 Ser Arg Ile Lys Arg Lys Lys Gln Glu Met Lys Lys Arg Lys Thr Arg  
 405                      410                      415  
 Gly Arg Cys Glu Val Val Ile Met Glu Asp Ala Pro Asp Tyr Tyr Ala  
 420                      425                      430  
 Val Glu Asp Ile Phe Ser Glu Ile Ser Asp Ile Asp Glu Thr Ile His  
 435                      440                      445  
 Asp Ile Lys Ile Ser Asp Phe Met Glu Thr Thr Asp Tyr Phe Glu Thr  
 450                      455                      460  
 Thr Asp Asn Glu Ile Thr Asp Ile Asn Glu Asn Ile Cys Asp Ser Glu  
 465                      470                      475                      480  
 Asn Pro Asp His Asn Glu Val Pro Asn Asn Glu Thr Thr Asp Asn Asn  
 485                      490                      495  
 Glu Ser Ala Asp Asp His Glu Thr Thr Asp Asn Asn Glu Ser Ala Asp  
 500                      505                      510  
 Asp Asn Asn Glu Asn Pro Glu Asp Asn Asn Lys Asn Thr Asp Asp Asn  
 515                      520                      525  
 Glu Glu Asn Pro Asn Asn Asn Glu Asn Thr Tyr Gly Asn Asn Phe Phe  
 530                      535                      540  
 Lys Gly Gly Phe Trp Gly Ser His Gly Asn Asn Gln Asp Ser Ser Asp  
 545                      550                      555                      560  
 Ser Asp Asn Glu Ala Asp Glu Ala Ser Asp Asp Glu Asp Asn Asp Gly  
 565                      570                      575  
 Asn Glu Gly Asp Asn Glu Gly Ser Asp Asp Asp Gly Asn Glu Gly Asp  
 580                      585                      590  
 Asn Glu Gly Ser Asp Asp Asp Asp Arg Asp Ile Glu Tyr Tyr Glu Lys  
 595                      600                      605  
 Val Ile Glu Asp Phe Asp Lys Asp Gln Ala Asp Tyr Glu Asp Val Ile  
 610                      615                      620  
 Glu Ile Ile Ser Asp Glu Ser Val Glu Glu Glu Gly Ile Glu Glu Gly

625

630

635

640

Ile Gln Gln Asp Glu Asp Ile Tyr Glu Glu Gly Asn Tyr Glu Glu Glu  
645 650 655

Gly Ser Glu Asp Val Trp Glu Glu Gly Glu Asp Ser Asp Asp Ser Asp  
660 665 670

Leu Glu Asp Val Leu Gln Val Pro Asn Gly Trp Ala Asn Pro Gly Lys  
675 680 685

Arg Gly Lys Thr Gly  
690

&lt;210&gt; 7

&lt;211&gt; 2632

&lt;212&gt; DNA

&lt;213&gt; NM\_016463 HSPC195

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (931)..(1611)

&lt;223&gt;

&lt;400&gt; 7

ggcggctgag cctgagcggg gatgtagagg cggcggcagc agaggcggca ctggcggcaa 60  
gagcagacgc ccgagccgag cgagaagagc ggcagagcct tatccctga agccggggccc 120  
cgcgtccag cctgcccag ccgcgccca gccatgcgcg ccgcctgctg agtcggggcg 180  
ccgcacgctg agccctccgc ccgcgagccg cgctcagctc gggggtgatt agttgctttt 240  
tgttggtttt taatttgggc cgcggggagg gggaggaggg gcaggtgctg caggetcccc 300  
cccctccccg cctcggggcca gccgcggcgg cgcgactcgg gctccggacc cgggcactgc 360  
tggcggctgg agcggagcgc accgcggcgg tgggtgccag agcggagcgc agtccctgc 420  
cccgccttc cccctcgcc tcgcggcgac ggcggcgggtg gcggcttga cgactcggag 480  
agccgagtga agacatttcc acctggacac ctgacctgt gctgccctg agcagcgagg 540  
cccaccaggc atctctgttg tgggcagcag ggccaggctc tggctctgtg accctcggca 600  
gttggcaggc tccctctgca gtggggtctg ggcctcggcc ccacctgtc gagcctcggc 660  
ggtggctccc aggatgccgg cggcagtagc agcagcagca ccaatggcag cggtggcagt 720  
ggcagcagtg gcccaaaggc aggagcagca gacaagagt cagtgggtggc tgccgccgca 780  
ccagcctcag tggcagatga cacaccaccc ccgcagcgtc ggaacaagag cggtatcatc 840

agt	gag	cccc	tca	aca	aag	ag	cct	gcg	ccgc	tccc	gccgc	tctccc	acta	ctcttctttt	900		
ggc	agc	cagt	gtg	gtag	tgg	cggt	ggc	agc	atg	atg	ggc	gga	gag	tct	gct	gac	954
									Met	Met	Gly	Gly	Glu	Ser	Ala	Asp	
									1				5				
aag	gcc	act	gcg	gct	gca	gcc	gct	gcc	tcc	ctg	ttg	gcc	aat	ggg	cat		1002
Lys	Ala	Thr	Ala	Ala	Ala	Ala	Ala	Ala	Ser	Leu	Leu	Ala	Asn	Gly	His		
	10					15					20						
gac	ctg	gcg	gcg	gcc	atg	gcg	gtg	gac	aaa	agc	aac	cct	acc	tca	aag		1050
Asp	Leu	Ala	Ala	Ala	Met	Ala	Val	Asp	Lys	Ser	Asn	Pro	Thr	Ser	Lys		
25					30					35					40		
cac	aaa	agt	ggt	gct	gtg	gcc	agc	ctg	ctg	agc	aag	gca	gag	cgg	gcc		1098
His	Lys	Ser	Gly	Ala	Val	Ala	Ser	Leu	Leu	Ser	Lys	Ala	Glu	Arg	Ala		
			45						50					55			
acg	gag	ctg	gca	gcc	gag	gga	cag	ctg	acg	ctg	cag	cag	ttt	gcg	cag		1146
Thr	Glu	Leu	Ala	Ala	Glu	Gly	Gln	Leu	Thr	Leu	Gln	Gln	Phe	Ala	Gln		
			60					65					70				
tcc	aca	gag	atg	ctg	aag	cg	gtg	gtg	cag	gag	cat	ctc	ccg	ctg	atg		1194
Ser	Thr	Glu	Met	Leu	Lys	Arg	Val	Val	Gln	Glu	His	Leu	Pro	Leu	Met		
		75					80					85					
agc	gag	gcg	ggt	gct	ggc	ctg	cct	gac	atg	gag	gct	gtg	gca	ggt	gcc		1242
Ser	Glu	Ala	Gly	Ala	Gly	Leu	Pro	Asp	Met	Glu	Ala	Val	Ala	Gly	Ala		
		90				95					100						
gaa	gcc	ctc	aat	ggc	cag	tcc	gac	ttc	ccc	tac	ctg	ggc	gct	ttc	ccc		1290
Glu	Ala	Leu	Asn	Gly	Gln	Ser	Asp	Phe	Pro	Tyr	Leu	Gly	Ala	Phe	Pro		
105					110					115					120		
atc	aac	cca	ggc	ctc	ttc	att	atg	acc	ccg	gca	ggt	gtg	ttc	ctg	gcc		1338
Ile	Asn	Pro	Gly	Leu	Phe	Ile	Met	Thr	Pro	Ala	Gly	Val	Phe	Leu	Ala		
				125					130					135			
gag	agc	gcg	ctg	cac	atg	gcg	ggc	ctg	gct	gag	tac	ccc	atg	cag	gga		1386
Glu	Ser	Ala	Leu	His	Met	Ala	Gly	Leu	Ala	Glu	Tyr	Pro	Met	Gln	Gly		
			140					145					150				
gag	ctg	gcc	tct	gcc	atc	agc	tcc	ggc	aag	aag	aag	cgg	aaa	cg	tgc		1434
Glu	Leu	Ala	Ser	Ala	Ile	Ser	Ser	Gly	Lys	Lys	Lys	Arg	Lys	Arg	Cys		
		155					160					165					
ggc	atg	tgc	gcg	ccc	tgc	cg	cg	cgc	atc	aac	tgc	gag	cag	tgc	agc		1482
Gly	Met	Cys	Ala	Pro	Cys	Arg	Arg	Arg	Ile	Asn	Cys	Glu	Gln	Cys	Ser		
	170					175					180						
agt	tgt	agg	aat	cga	aag	act											

```

aagaaaaaaa tttaaactgc tttttcggaa gaacaacaac aaaaaagagg taaagacgaa 1871
tctataaagt accgagactt cctgggcaaa gaatggacaa tcagtttcct tctgtgtcgc 1931
atgtcgtatgt tgtctgtgca ggagatgcag tttttgtgta gagaatgtaa attttctgta 1991
accttttgaa atctagttac taataagcac tactgtaatt tagcacagtt taactccacc 2051
ctcatttaaa cttcctttga ttctttccga ccatgaaata gtgcatagtt tgcctggaga 2111
atccactcac gttcataaag agaatgttga tggcgccgtg tagaagccgc tctgtatcca 2171
tccacgcgtg cagagctgcc agcagggagc tcacagaagg ggagggagca ccaggccagc 2231
tgagctgcac ccacagtccc gagactggga tccccaccc caacagtgat tttggaaaaa 2291
aaaatgaaag ttctgttcgt ttatccattg cgatctgggg agcccatct cgatatttcc 2351
aatcctggct acttttctta gagaaaataa gtctttttt tctggccttg ctaatggcaa 2411
cagaagaaag ggcttctttg cgtggtcccc tgcgtggtgg ggtgggtccc cagggggccc 2471
cctgcgccct ggccccccct gccacggcc agcttcctgc tgatgaacat gctgtttgta 2531
ttgttttagg aaaccaggct gttttgtgaa taaaacgaat gcattgttgt gtcacgaaaa 2591
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa a 2632

```

&lt;210&gt; 8

&lt;211&gt; 227

&lt;212&gt; PRT

&lt;213&gt; NM\_016463 HSPC195

&lt;400&gt; 8

```

Met Met Gly Gly Glu Ser Ala Asp Lys Ala Thr Ala Ala Ala Ala Ala
1          5          10          15

Ala Ser Leu Leu Ala Asn Gly His Asp Leu Ala Ala Ala Met Ala Val
20          25          30

Asp Lys Ser Asn Pro Thr Ser Lys His Lys Ser Gly Ala Val Ala Ser
35          40          45

Leu Leu Ser Lys Ala Glu Arg Ala Thr Glu Leu Ala Ala Glu Gly Gln
50          55          60

Leu Thr Leu Gln Gln Phe Ala Gln Ser Thr Glu Met Leu Lys Arg Val
65          70          75          80

Val Gln Glu His Leu Pro Leu Met Ser Glu Ala Gly Ala Gly Leu Pro
85          90          95

Asp Met Glu Ala Val Ala Gly Ala Glu Ala Leu Asn Gly Gln Ser Asp
100         105         110

```

Phe Pro Tyr Leu Gly Ala Phe Pro Ile Asn Pro Gly Leu Phe Ile Met  
115 120 125

Thr Pro Ala Gly Val Phe Leu Ala Glu Ser Ala Leu His Met Ala Gly  
130 135 140

Leu Ala Glu Tyr Pro Met Gln Gly Glu Leu Ala Ser Ala Ile Ser Ser  
145 150 155 160

Gly Lys Lys Lys Arg Lys Arg Cys Gly Met Cys Ala Pro Cys Arg Arg  
165 170 175

Arg Ile Asn Cys Glu Gln Cys Ser Ser Cys Arg Asn Arg Lys Thr Gly  
180 185 190

His Gln Ile Cys Lys Phe Arg Lys Cys Glu Glu Leu Lys Lys Lys Pro  
195 200 205

Ser Ala Ala Leu Glu Lys Val Met Leu Pro Thr Gly Ala Ala Phe Arg  
210 215 220

Trp Phe Gln  
225

<210> 9

<211> 2140

<212> DNA

<213> NM\_017697 FLJ20171

<220>

<221> CDS

<222> (58)..(1131)

<223>

<400> 9

gaattcaaga aatgttgccc tggttcacct gatattgaca aactggacgt tgccaca 57

atg aca gag tat tta aat ttt gag aag agt agt tca gtc tct cga tat 105  
Met Thr Glu Tyr Leu Asn Phe Glu Lys Ser Ser Val Ser Arg Tyr  
1 5 10 15

gga gcc tct caa gtt gaa gat atg ggg aat ata att tta gca atg att 153  
Gly Ala Ser Gln Val Glu Asp Met Gly Asn Ile Ile Leu Ala Met Ile  
20 25 30

tca gag cct tat aat cac agg ttt tca gat cca gag aga gtg aat tac 201  
Ser Glu Pro Tyr Asn His Arg Phe Ser Asp Pro Glu Arg Val Asn Tyr  
35 40 45

aag ttt gaa agt gga act tgc agc aag atg gaa ctt att gat gat aac Lys Phe Glu Ser Gly Thr Cys Ser Lys Met Glu Leu Ile Asp Asp Asn 50 55 60	249
acc gta gtc agg gca cga ggt tta cca tgg cag tct tca gat caa gat Thr Val Val Arg Ala Arg Gly Leu Pro Trp Gln Ser Ser Asp Gln Asp 65 70 75 80	297
att gca aga ttc ttc aaa gga ctc aat att gcc aag gga ggt gca gca Ile Ala Arg Phe Phe Lys Gly Leu Asn Ile Ala Lys Gly Gly Ala Ala 85 90 95	345
ctt tgt ctg aat gct cag ggt cga agg aac gga gaa gct ctg gtt agg Leu Cys Leu Asn Ala Gln Gly Arg Arg Asn Gly Glu Ala Leu Val Arg 100 105 110	393
ttt gta agt gag gag cac cga gac cta gca cta cag agg cac aaa cat Phe Val Ser Glu Glu His Arg Asp Leu Ala Leu Gln Arg His Lys His 115 120 125	441
cac atg ggg acc cgg tat att gag gtt tac aaa gca aca ggt gaa gat His Met Gly Thr Arg Tyr Ile Glu Val Tyr Lys Ala Thr Gly Glu Asp 130 135 140	489
ttc ctt aaa att gct ggt ggt act tcc aat gag gta gcc cag ttt ctc Phe Leu Lys Ile Ala Gly Gly Thr Ser Asn Glu Val Ala Gln Phe Leu 145 150 155 160	537
tcc aag gaa aat caa gtc att gtt cgc atg cgg ggg ctc cct ttc acg Ser Lys Glu Asn Gln Val Ile Val Arg Met Arg Gly Leu Pro Phe Thr 165 170 175	585
gcc aca gct gaa gaa gtg gtg gcc ttc ttt gga cag cat tgc cct att Ala Thr Ala Glu Glu Val Val Ala Phe Phe Gly Gln His Cys Pro Ile 180 185 190	633
act ggg gga aag gaa ggc atc ctc ttt gtc acc tac cca gat ggt agg Thr Gly Gly Lys Glu Gly Ile Leu Phe Val Thr Tyr Pro Asp Gly Arg 195 200 205	681
cca aca ggg gac gct ttt gtc ctc ttt gcc tgt gag gaa tat gca cag Pro Thr Gly Asp Ala Phe Val Leu Phe Ala Cys Glu Glu Tyr Ala Gln 210 215 220	729
aat gcg ttg agg aag cat aaa gac ttg ttg ggt aaa aga tac att gaa Asn Ala Leu Arg Lys His Lys Asp Leu Leu Gly Lys Arg Tyr Ile Glu 225 230 235 240	777
ctc ttc agg agc aca gca gct gaa gtt cag cag gtg ctg aat cga ttc Leu Phe Arg Ser Thr Ala Ala Glu Val Gln Gln Val Leu Asn Arg Phe 245 250 255	825
tcc tcg gcc cct ctc att cca ctt cca acc cct ccc att att cca gta Ser Ser Ala Pro Leu Ile Pro Leu Pro Thr Pro Pro Ile Ile Pro Val 260 265 270	873
cta cct cag caa ttt gtg ccc cct aca aat gtt aga gac tgt ata cgc Leu Pro Gln Gln Phe Val Pro Pro Thr Asn Val Arg Asp Cys Ile Arg 275 280 285	921
ctt cga ggt ctt ccc tat gca gcc aca att gag gac atc ctg gat ttc Leu Arg Gly Leu Pro Tyr Ala Ala Thr Ile Glu Asp Ile Leu Asp Phe 290 295 300	969
ctg ggg gag ttc gcc aca gat att cgt act cat ggg gtt cac atg gtt Leu Gly Glu Phe Ala Thr Asp Ile Arg Thr His Gly Val His Met Val 305 310 315 320	1017



ttg aat cac cag ggc cgc cca tca gga gat gcc ttt atc cag atg aag 1065  
 Leu Asn His Gln Gly Arg Pro Ser Gly Asp Ala Phe Ile Gln Met Lys  
 325 330 335  
 tct gcg gac aga gca ttt atg gct gca cag aag tgt cat aaa aaa aaa 1113  
 Ser Ala Asp Arg Ala Phe Met Ala Ala Gln Lys Cys His Lys Lys Lys  
 340 345 350  
 cat gaa gga cag ata tgt tgaagtcttt cagtgttcag ctgaggagat 1161  
 His Glu Gly Gln Ile Cys  
 355  
 gaactttgtg ttaatggggg gcactttaaa tcgaaatggc ttatccccac cgccatgtaa 1221  
 gttaccatgc ctgtctcctc cctcctacac atttcagct cctgctgcag ttattcctac 1281  
 agaagctgcc atttaccagc cctctgtgat tttgaatcca cgagcactgc agccctccac 1341  
 agcgtactac ccagcaggca ctcagctctt catgaactac acagcgtact atcccagccc 1401  
 cccagggttcg cctaatagtc ttggctactt ccctacagct gctaacttta gcggtgtccc 1461  
 tccacagcct ggcacggtgg tcagaatgca gggcctggcc tacaatactg gagttaagga 1521  
 aattcttaac ttcttccaag gttaccagt tttgaaagat gtatggtgat cttgaaacct 1581  
 ccagacacaa gaaaacttct agcaaattca ggggaagttt gtctacactc aggctgcagt 1641  
 attttcagca aacttgattg gacaaacggg cctgtgcctt atcttttggg ggagtgaaaa 1701  
 agtttgagct agtgaagcca aatcgtaact tacagcaagc agcatgcagc atacctggct 1761  
 ctgttgctgat tgcaaatagg catttaaaat gtgaatttgg aatcagatgt ctccattact 1821  
 tccagttaaa gtggcatcat aggtgtttcc taagttttaa gtcttggata aaaactccac 1881  
 cagtgtctac catctccacc atgaactctg ttaaggaagc ttcatttttg tatattcccg 1941  
 ctcttttctc ttcatttccc tgtcttctgc ataatcatgc cttcttgcta agtaattcaa 2001  
 gcataagatc ttggaataat aaaatcacia tcttaggaga aagaataaaa ttgttatatt 2061  
 cccagtctct tggccatgat gatattctat gattaaaaac aaattaaatt ttaaacacc 2121  
 tgaaaaaaaa aaaaaaaaaa 2140

&lt;210&gt; 10

&lt;211&gt; 358

&lt;212&gt; PRT

&lt;213&gt; NM\_017697 FLJ20171

&lt;400&gt; 10

Met Thr Glu Tyr Leu Asn Phe Glu Lys Ser Ser Ser Val Ser Arg Tyr  
 1 5 10 15

Gly Ala Ser Gln Val Glu Asp Met Gly Asn Ile Ile Leu Ala Met Ile  
 20 25 30

Ser Glu Pro Tyr Asn His Arg Phe Ser Asp Pro Glu Arg Val Asn Tyr  
35 40 45

Lys Phe Glu Ser Gly Thr Cys Ser Lys Met Glu Leu Ile Asp Asp Asn  
50 55 60

Thr Val Val Arg Ala Arg Gly Leu Pro Trp Gln Ser Ser Asp Gln Asp  
65 70 75 80

Ile Ala Arg Phe Phe Lys Gly Leu Asn Ile Ala Lys Gly Gly Ala Ala  
85 90 95

Leu Cys Leu Asn Ala Gln Gly Arg Arg Asn Gly Glu Ala Leu Val Arg  
100 105 110

Phe Val Ser Glu Glu His Arg Asp Leu Ala Leu Gln Arg His Lys His  
115 120 125

His Met Gly Thr Arg Tyr Ile Glu Val Tyr Lys Ala Thr Gly Glu Asp  
130 135 140

Phe Leu Lys Ile Ala Gly Gly Thr Ser Asn Glu Val Ala Gln Phe Leu  
145 150 155 160

Ser Lys Glu Asn Gln Val Ile Val Arg Met Arg Gly Leu Pro Phe Thr  
165 170 175

Ala Thr Ala Glu Glu Val Val Ala Phe Phe Gly Gln His Cys Pro Ile  
180 185 190

Thr Gly Gly Lys Glu Gly Ile Leu Phe Val Thr Tyr Pro Asp Gly Arg  
195 200 205

Pro Thr Gly Asp Ala Phe Val Leu Phe Ala Cys Glu Glu Tyr Ala Gln  
210 215 220

Asn Ala Leu Arg Lys His Lys Asp Leu Leu Gly Lys Arg Tyr Ile Glu  
225 230 235 240

Leu Phe Arg Ser Thr Ala Ala Glu Val Gln Gln Val Leu Asn Arg Phe  
245 250 255

Ser Ser Ala Pro Leu Ile Pro Leu Pro Thr Pro Pro Ile Ile Pro Val  
260 265 270

Leu Pro Gln Gln Phe Val Pro Pro Thr Asn Val Arg Asp Cys Ile Arg  
275 280 285

Leu Arg Gly Leu Pro Tyr Ala Ala Thr Ile Glu Asp Ile Leu Asp Phe  
290 295 300

Leu Gly Glu Phe Ala Thr Asp Ile Arg Thr His Gly Val His Met Val  
305 310 315 320

Leu Asn His Gln Gly Arg Pro Ser Gly Asp Ala Phe Ile Gln Met Lys  
325 330 335

Ser Ala Asp Arg Ala Phe Met Ala Ala Gln Lys Cys His Lys Lys Lys  
340 345 350

His Glu Gly Gln Ile Cys  
355

<210> 11

<211> 2808

<212> DNA

<213> NM\_052886 MAL2

<220>

<221> CDS

<222> (80)..(607)

<223>

<400> 11

ggcggcggcg gcaggagccc gggaggcgga ggcgggagggc ggcggcggcg cgcggagacg 60

cagcagcggc agcggcagc atg tcg gcc ggc gga gcg tca gtc ccg ccg ccc 112  
Met Ser Ala Gly Gly Ala Ser Val Pro Pro Pro  
1 5 10

ccg aac ccc gcc gtg tcc ttc ccg ccg ccc cgg gtc acc ctg ccc gcc 160  
Pro Asn Pro Ala Val Ser Phe Pro Pro Pro Arg Val Thr Leu Pro Ala  
15 20 25

ggc ccc gac atc ctg cgg acc tac tcg ggc gcc ttc gtc tgc ctg gag 208  
Gly Pro Asp Ile Leu Arg Thr Tyr Ser Gly Ala Phe Val Cys Leu Glu  
30 35 40

att ctg ttc ggg ggt ctt gtc tgg att ttg gtt gcc tcc tcc aat gtt 256  
Ile Leu Phe Gly Gly Leu Val Trp Ile Leu Val Ala Ser Ser Asn Val  
45 50 55

cct cta cct cta cta caa gga tgg gtc atg ttt gtg tcc gtg aca gcg 304  
Pro Leu Pro Leu Leu Gln Gly Trp Val Met Phe Val Ser Val Thr Ala  
60 65 70 75

ttt ttc ttt tcg ctc ctc ttt ctg ggc atg ttc ctc tct ggc atg gtg 352  
Phe Phe Phe Ser Leu Leu Phe Leu Gly Met Phe Leu Ser Gly Met Val  
80 85 90

gct caa att gat gct aac tgg aac ttc ctg gat ttt gcc tac cat ttt 400  
Ala Gln Ile Asp Ala Asn Trp Asn Phe Leu Asp Phe Ala Tyr His Phe  
95 100 105

aca gta ttt gtc ttc tat ttt gga gcc ttt tta ttg gaa gca gca gcc 448  
 Thr Val Phe Val Phe Tyr Phe Gly Ala Phe Leu Leu Glu Ala Ala Ala  
 110 115 120

aca tcc ctg cat gat ttg cat tgc aat aca acc ata acc ggg cag cca 496  
 Thr Ser Leu His Asp Leu His Cys Asn Thr Thr Ile Thr Gly Gln Pro  
 125 130 135

ctg ctg agt gat aac cag tat aac ata aac gta gca gcc tca att ttt 544  
 Leu Leu Ser Asp Asn Gln Tyr Asn Ile Asn Val Ala Ala Ser Ile Phe  
 140 145 150 155

gcc ttt atg acg aca gct tgt tat ggt tgc agt ttg ggt ctg gct tta 592  
 Ala Phe Met Thr Thr Ala Cys Tyr Gly Cys Ser Leu Gly Leu Ala Leu  
 160 165 170

cga aga tgg cga ccg taacactcct tagaaactgg cagtcgtatg ttagtttcac 647  
 Arg Arg Trp Arg Pro  
 175

ttgtctactt tatatgtctg atcaatttgg ataccatttt gtccagatgc aaaaacattc 707

caaaagtaat gtgttttagta gagagagact ctaagctcaa gttctggttt atttcattga 767

tggaatgtta attttattat gatattaaag aaatggcctt ttattttaca tctctcccct 827

ttttcccttt ccccttttat tttctcctt tttcttctga aagtttccctt ttatgtccat 887

aaaatacaaa tatattgttc ataaaaaatt agtatccctt ttgtttgggt gctgagtcac 947

ctgaacctta attttaattg gtaattacag cccttaaaaa aaacacattt caaataggct 1007

tcccactaaa ctctatatat tagtgtaaac caggaattgg cacacttttt ttagaatggg 1067

ccagatggta aatatttatg cttcacgggc catacagtct ctgtcacaac tattcagttc 1127

tgctagtata gcgtgaaagc agctatacac aatacagaaa tgaatgagtg tggttatgtt 1187

ctaataaaac ttatttataa aaacaagggg aggctgggtt tagcctgtgg gccatagttt 1247

gtcaaccact ggtgtaaaac cttagttata tatgatctgc attttcttga actgatcatt 1307

gaaaacttat aaacctaaca gaaaagccac ataatatita gtgtcattat gcaataatca 1367

cattgccttt gtgttaaatg tcaataactt acctttggag aatacttacc tttggaggaa 1427

tgtataaaat ttctcaggca gagtcctgga tataggaaaa agtaatttat gaagtaaact 1487

tcagttgctt aatcaaaact atgatagtct aacaactgag caagatcctc atctgagagt 1547

gcttaaaatg ggatccccag agaccattaa ccaatactgg aactggtatc tagctactga 1607

tgtcttactt tgagtttatt tatgcttcag aatacagttg tttgccctgt gcatgaatat 1667

acccatatth gtgtgtggat atgtgaagct tttccaaata gagctctcag aagaattaag 1727

tttttacttc taattattht gcattacttt gagttaaatt tgaatagagt attaaatata 1787

aagttgtaga ttcttatgtg tttttgtatt agcccagaca tctgtaagt ttttgcaactg 1847

gtgacagaca aaatctgttt taaaatcata tccagcacia aaactatttc tggctgaata 1907

gcacagaaaa gtattttaac ctacctgtag agatcctcgt catggaaagg tgccaaactg 1967

ttttgaatgg aaggacaagt aagagtgagg ccacagttcc caccacacga gggcttttgt 2027

attgttctac tttttcagcc ctttactttc tggctgaagc atccccttgg agtgccatgt 2087

ataagttggg ctattagagt tcatggaaca tagaacaacc atgaatgagt ggcattgatcc 2147  
 gtgcttaatg atcaagtgtt acttatctaa taatcctcta gaaagaaccc tgttagatct 2207  
 tggtttgtga taaaaatata aagacagaag acatgaggaa aaacaaaagg tttagaggaaa 2267  
 tcaggcatat gactttatac ttaacatcag atcttttcta taatataccta ctacttttgt 2327  
 tttcctagct ccataccaca cacctaaacc tgtattatga attacatatt acaaagtcatt 2387  
 aaatgtgccca tatggatata cagtacattc tagttggaat cgtttactct gctagaattt 2447  
 aggtgtgaga ttttttgttt ccaggtata gcaggcttat gtttggtggc attaaattgg 2507  
 tttctttaa atgcttttgt ggcacttttg taaacagatt gcttctagat tgttacaac 2567  
 caagcctaag acacatctgt gaatacttag attttagct taatcacatt ctgacttgt 2627  
 gagttgaatg acaaacgagt tgaacaaaaa ttatggcatt taagaattta acatgtctta 2687  
 gctgtaaaaa tgagaaagtg ttggttggtt ttaaaatctg gtaactccat gatgaaaaga 2747  
 aatttatttt atacgtgtta tgtctctaataaagtattca ttgataaaaa aaaaaaaaaa 2807  
 a 2808

<210> 12

<211> 176

<212> PRT

<213> NM\_052886 MAL2

<400> 12

Met Ser Ala Gly Gly Ala Ser Val Pro Pro Pro Pro Asn Pro Ala Val  
1 5 10 15

Ser Phe Pro Pro Pro Arg Val Thr Leu Pro Ala Gly Pro Asp Ile Leu  
20 25 30

Arg Thr Tyr Ser Gly Ala Phe Val Cys Leu Glu Ile Leu Phe Gly Gly  
35 40 45

Leu Val Trp Ile Leu Val Ala Ser Ser Asn Val Pro Leu Pro Leu Leu  
50 55 60

Gln Gly Trp Val Met Phe Val Ser Val Thr Ala Phe Phe Phe Ser Leu  
65 70 75 80

Leu Phe Leu Gly Met Phe Leu Ser Gly Met Val Ala Gln Ile Asp Ala  
85 90 95

Asn Trp Asn Phe Leu Asp Phe Ala Tyr His Phe Thr Val Phe Val Phe  
100 105 110

Tyr Phe Gly Ala Phe Leu Leu Glu Ala Ala Ala Thr Ser Leu His Asp

115 120 125

Leu His Cys Asn Thr Thr Ile Thr Gly Gln Pro Leu Leu Ser Asp Asn  
130 135 140

Gln Tyr Asn Ile Asn Val Ala Ala Ser Ile Phe Ala Phe Met Thr Thr  
145 150 155 160

Ala Cys Tyr Gly Cys Ser Leu Gly Leu Ala Leu Arg Arg Trp Arg Pro  
165 170 175

<210> 13  
<211> 4171  
<212> DNA  
<213> NM\_015238 KIAA0869, KIBRA

<220>  
<221> CDS  
<222> (67)..(3405)  
<223>

<400> 13  
ctggagccgc tgagcccccg ctgcggccgg gagctgcatg ggggagcgcc ggcagcgctt 60  
gggaag atg ccc cgg ccg gag ctg ccc ctg ccg gag ggc tgg gag gag 108  
Met Pro Arg Pro Glu Leu Pro Leu Pro Glu Gly Trp Glu Glu  
1 5 10  
gcg cgc gac ttc gac ggc aag gtc tac tac ata gac cac acg aac cgc 156  
Ala Arg Asp Phe Asp Gly Lys Val Tyr Tyr Ile Asp His Thr Asn Arg  
15 20 25 30  
acc acc agc tgg atc gac ccg ccg gag agg tac acc aaa ccg ctc acc 204  
Thr Thr Ser Trp Ile Asp Pro Arg Asp Arg Tyr Thr Lys Pro Leu Thr  
35 40 45  
ttt gct gac tgc att agt gat gag ttg ccg cta gga tgg gaa gag gca 252  
Phe Ala Asp Cys Ile Ser Asp Glu Leu Pro Leu Gly Trp Glu Glu Ala  
50 55 60  
tat gac cca cag gtt gga gat tac ttc ata gac cac aac acc aaa acc 300  
Tyr Asp Pro Gln Val Gly Asp Tyr Phe Ile Asp His Asn Thr Lys Thr  
65 70 75  
act cag att gag gat cct cga gta caa tgg ccg ccg gag cag gaa cat 348  
Thr Gln Ile Glu Asp Pro Arg Val Gln Trp Arg Arg Glu Gln Glu His  
80 85 90  
atg ctg aag gat tac ctg gtg gtg gcc cag gag gct ctg agt gca caa 396  
Met Leu Lys Asp Tyr Leu Val Val Ala Gln Glu Ala Leu Ser Ala Gln  
95 100 105 110  
aag gag atc tac cag gtg aag cag cag cgc ctg gag ctt gca cag cag 444

Lys	Glu	Ile	Tyr	Gln	Val	Lys	Gln	Gln	Arg	Leu	Glu	Leu	Ala	Gln	Gln		
				115					120					125			
gag	tac	cag	caa	ctg	cat	gcc	gtc	tgg	gag	cat	aag	ctg	ggc	tcc	cag		492
Glu	Tyr	Gln	Gln	Leu	His	Ala	Val	Trp	Glu	His	Lys	Leu	Gly	Ser	Gln		
			130					135					140				
gtc	agc	ttg	gtc	tct	ggt	tca	tca	tcc	agc	tcc	aag	tat	gac	cct	gag		540
Val	Ser	Leu	Val	Ser	Gly	Ser	Ser	Ser	Ser	Ser	Lys	Tyr	Asp	Pro	Glu		
		145					150					155					
atc	ctg	aaa	gct	gaa	att	gcc	act	gca	aaa	tcc	cgg	gtc	aac	aag	ctg		588
Ile	Leu	Lys	Ala	Glu	Ile	Ala	Thr	Ala	Lys	Ser	Arg	Val	Asn	Lys	Leu		
	160					165					170						
aag	aga	gag	atg	gtt	cac	ctc	cag	cac	gag	ctg	cag	ttc	aaa	gag	cgt		636
Lys	Arg	Glu	Met	Val	His	Leu	Gln	His	Glu	Leu	Gln	Phe	Lys	Glu	Arg		
175					180					185					190		
ggc	ttt	cag	acc	ctg	aag	aaa	atc	gat	aag	aaa	atg	tct	gat	gct	cag		684
Gly	Phe	Gln	Thr	Leu	Lys	Lys	Ile	Asp	Lys	Lys	Met	Ser	Asp	Ala	Gln		
				195					200					205			
ggc	agc	tac	aaa	ctg	gat	gaa	gct	cag	gct	gtc	ttg	aga	gaa	aca	aaa		732
Gly	Ser	Tyr	Lys	Leu	Asp	Glu	Ala	Gln	Ala	Val	Leu	Arg	Glu	Thr	Lys		
			210				215						220				
gcc	atc	aaa	aag	gct	att	acc	tgt	ggg	gaa	aag	gaa	aag	caa	gat	ctc		780
Ala	Ile	Lys	Lys	Ala	Ile	Thr	Cys	Gly	Glu	Lys	Glu	Lys	Gln	Asp	Leu		
	225						230					235					
att	aag	agc	ctt	gcc	atg	ttg	aag	gac	ggc	ttc	cgc	act	gac	agg	ggg		828
Ile	Lys	Ser	Leu	Ala	Met	Leu	Lys	Asp	Gly	Phe	Arg	Thr	Asp	Arg	Gly		
	240					245					250						
tct	cac	tca	gac	ctg	tgg	tcc	agc	agc	agc	tct	ctg	gag	agt	tcg	agt		876
Ser	His	Ser	Asp	Leu	Trp	Ser	Ser	Ser	Ser	Ser	Leu	Glu	Ser	Ser	Ser		
255					260					265					270		
ttc	ccg	cta	ccg	aaa	cag	tac	ctg	gat	gtg	agc	tcc	cag	aca	gac	atc		924
Phe	Pro	Leu	Pro	Lys	Gln	Tyr	Leu	Asp	Val	Ser	Ser	Gln	Thr	Asp	Ile		
				275					280					285			
tcg	gga	agc	ttc	ggc	atc	aac	agc	aac	aat	cag	ttg	gca	gag	aag	gtc		972
Ser	Gly	Ser	Phe	Gly	Ile	Asn	Ser	Asn	Asn	Gln	Leu	Ala	Glu	Lys	Val		
			290				295						300				
aga	ttg	cgc	ctt	cga	tat	gaa	gag	gct	aag	aga	agg	atc	gcc	aac	ctg		1020
Arg	Leu	Arg	Leu	Arg	Tyr	Glu	Glu	Ala	Lys	Arg	Arg	Ile	Ala	Asn	Leu		
			305				310					315					
aag	atc	cag	ctg	gcc	aag	ctt	gac	agt	gag	gcc	tgg	cct	ggg	gtg	ctg		1068
Lys	Ile	Gln	Leu	Ala	Lys	Leu	Asp	Ser	Glu	Ala	Trp	Pro	Gly	Val	Leu		
	320					325					330						
gac	tca	gag	agg	gac	cgc	ctg	atc	ctt	atc	aac	gag	aag	gag	gag	ctg		1116
Asp	Ser	Glu	Arg	Asp	Arg	Leu	Ile	Leu	Ile	Asn	Glu	Lys	Glu	Glu	Leu		
335				340					345					350			
ctg	aag	gag	atg	cgc	ttc	atc	agc	ccc	cgc	aag	tgg	acc	cag	ggg	gag		1164
Leu	Lys	Glu	Met	Arg	Phe	Ile	Ser	Pro	Arg	Lys	Trp	Thr	Gln	Gly	Glu		
				355					360					365			
gtg	gag	cag	ctg	gag	atg	gcc	cgg	aag	cgg	ctg	gaa	aag	gac	ctg	cag		1212
Val	Glu	Gln	Leu	Glu	Met	Ala	Arg	Lys	Arg	Leu	Glu	Lys	Asp	Leu	Gln		
			370				375						380				

gca gcc cgg gac acc cag agc aag gcg ctg acg gag agg tta aag tta Ala Ala Arg Asp Thr Gln Ser Lys Ala Leu Thr Glu Arg Leu Lys Leu 385 390 395	1260
aac agt aag agg aac cag ctt gtg aga gaa ctg gag gaa gcc acc cgg Asn Ser Lys Arg Asn Gln Leu Val Arg Glu Leu Glu Ala Thr Arg 400 405 410	1308
cag gtg gca act ctg cac tcc cag ctg aaa agt ctc tca agc agc atg Gln Val Ala Thr Leu His Ser Gln Leu Lys Ser Leu Ser Ser Met 415 420 425 430	1356
cag tcc ctg tcc tca ggc agc agc ccc gga tcc ctc acg tcc agc cgg Gln Ser Leu Ser Ser Gly Ser Ser Pro Gly Ser Leu Thr Ser Ser Arg 435 440 445	1404
ggc tcc ctg gtt gca tcc agc ctg gac tcc tcc act tca gcc agc ttc Gly Ser Leu Val Ala Ser Ser Leu Asp Ser Ser Thr Ser Ala Ser Phe 450 455 460	1452
act gac ctc tac tat gac ccc ttt gag cag ctg gac tca gag ctg cag Thr Asp Leu Tyr Tyr Asp Pro Phe Glu Gln Leu Asp Ser Glu Leu Gln 465 470 475	1500
agc aag gtg gag ttc ctg ctc ctg gag ggg gcc acc ggc ttc cgg ccc Ser Lys Val Glu Phe Leu Leu Leu Glu Gly Ala Thr Gly Phe Arg Pro 480 485 490	1548
tca ggc tgc atc acc acc atc cac gag gat gag gtg gcc aag acc cag Ser Gly Cys Ile Thr Thr Ile His Glu Asp Glu Val Ala Lys Thr Gln 495 500 505 510	1596
aag gca gag gga ggt ggc cgc ctg cag gct ctg cgt tcc ctg tct ggc Lys Ala Glu Gly Gly Arg Leu Gln Ala Leu Arg Ser Leu Ser Gly 515 520 525	1644
acc cca aag tcc atg acc tcc cta tcc cca cgt tcc tct ctc tcc tcc Thr Pro Lys Ser Met Thr Ser Leu Ser Pro Arg Ser Ser Leu Ser Ser 530 535 540	1692
ccc tcc cca ccc tgt tcc cct ctc atg gct gac ccc ctc ctg gct ggt Pro Ser Pro Pro Cys Ser Pro Leu Met Ala Asp Pro Leu Leu Ala Gly 545 550 555	1740
gat gcc ttc ctc aac tcc ttg gag ttt gaa gac ccg gag ctg agt gcc Asp Ala Phe Leu Asn Ser Leu Glu Phe Glu Asp Pro Glu Leu Ser Ala 560 565 570	1788
act ctt tgt gaa ctg agc ctt ggt aac agc gcc cag gaa aga tac cgg Thr Leu Cys Glu Leu Ser Leu Gly Asn Ser Ala Gln Glu Arg Tyr Arg 575 580 585 590	1836
ctg gag gaa cca gga acg gag ggc aag cag ctg ggc caa gct gtg aat Leu Glu Glu Pro Gly Thr Glu Gly Lys Gln Leu Gly Gln Ala Val Asn 595 600 605	1884
acg gcc cag ggg tgt ggc ctg aaa gtg gcc tgt gtc tca gcc gcc gta Thr Ala Gln Gly Cys Gly Leu Lys Val Ala Cys Val Ser Ala Ala Val 610 615 620	1932
tcg gac gag tca gtg gct gga gac agt ggt gtg tac gag gct tcc gtg Ser Asp Glu Ser Val Ala Gly Asp Ser Gly Val Tyr Glu Ala Ser Val 625 630 635	1980
cag aga ctg ggt gct tca gaa gct gct gca ttt gac agt gac gaa tcg	2028



Gln	Arg	Leu	Gly	Ala	Ser	Glu	Ala	Ala	Ala	Phe	Asp	Ser	Asp	Glu	Ser		
640						645					650						
gaa	gca	gtg	ggg	gcg	acc	cga	att	cag	att	gcc	ctg	aag	tat	gat	gag	2076	
Glu	Ala	Val	Gly	Ala	Thr	Arg	Ile	Gln	Ile	Ala	Leu	Lys	Tyr	Asp	Glu		
655					660					665					670		
aag	aat	aag	caa	ttt	gca	ata	tta	atc	atc	cag	ctg	agt	aac	ctt	tct	2124	
Lys	Asn	Lys	Gln	Phe	Ala	Ile	Leu	Ile	Ile	Gln	Leu	Ser	Asn	Leu	Ser		
				675						680					685		
gct	ctg	ttg	cag	caa	caa	gac	cag	aaa	gtg	aat	atc	cgc	gtg	gct	gtc	2172	
Ala	Leu	Leu	Gln	Gln	Gln	Asp	Gln	Lys	Val	Asn	Ile	Arg	Val	Ala	Val		
				690					695					700			
ctt	cct	tgc	tct	gaa	agc	aca	acc	tgc	ctg	ttc	cgg	acc	cgg	cct	ctg	2220	
Leu	Pro	Cys	Ser	Glu	Ser	Thr	Thr	Cys	Leu	Phe	Arg	Thr	Arg	Pro	Leu		
		705						710				715					
gac	gcc	tca	gac	act	cta	gtg	ttc	aat	gag	gtg	ttc	tgg	gta	tcc	atg	2268	
Asp	Ala	Ser	Asp	Thr	Leu	Val	Phe	Asn	Glu	Val	Phe	Trp	Val	Ser	Met		
		720				725					730						
tcc	tat	cca	gcc	ctt	cac	cag	aag	acc	tta	aga	gtc	gat	gtc	tgt	acc	2316	
Ser	Tyr	Pro	Ala	Leu	His	Gln	Lys	Thr	Leu	Arg	Val	Asp	Val	Cys	Thr		
735					740					745					750		
acc	gac	agg	agc	cat	ctg	gaa	gag	tgc	ctg	gga	ggc	gcc	cag	atc	agc	2364	
Thr	Asp	Arg	Ser	His	Leu	Glu	Glu	Cys	Leu	Gly	Gly	Ala	Gln	Ile	Ser		
				755					760						765		
ctg	gcg	gag	gtc	tgc	cgg	tct	ggg	gag	agg	tcg	act	cgc	tgg	tac	aac	2412	
Leu	Ala	Glu	Val	Cys	Arg	Ser	Gly	Glu	Arg	Ser	Thr	Arg	Trp	Tyr	Asn		
			770					775						780			
ctt	ctc	agc	tac	aaa	tac	ttg	aag	aag	cag	agc	agg	gag	ctc	aag	cca	2460	
Leu	Leu	Ser	Tyr	Lys	Tyr	Leu	Lys	Lys	Gln	Ser	Arg	Glu	Leu	Lys	Pro		
		785						790					795				
gtg	gga	gtt	atg	gcc	cct	gcc	tca	ggg	cct	gcc	agc	acg	gac	gct	gtg	2508	
Val	Gly	Val	Met	Ala	Pro	Ala	Ser	Gly	Pro	Ala	Ser	Thr	Asp	Ala	Val		
		800				805						810					
tct	gct	ctg	ttg	gaa	cag	aca	gca	gtg	gag	ctg	gag	aag	agg	cag	gag	2556	
Ser	Ala	Leu	Leu	Glu	Gln	Thr	Ala	Val	Glu	Leu	Glu	Lys	Arg	Gln	Glu		
815					820					825					830		
ggc	agg	agc	agc	aca	cag	aca	ctg	gaa	gac	agc	tgg	agg	tat	gag	gag	2604	
Gly	Arg	Ser	Ser	Thr	Gln	Thr	Leu	Glu	Asp	Ser	Trp	Arg	Tyr	Glu	Glu		
				835					840						845		
acc	agt	gag	aat	gag	gca	gta	gcc	gag	gaa	gag	gag	gag	gag	gtg	gag	2652	
Thr	Ser	Glu	Asn	Glu	Ala	Val	Ala	Glu	Glu	Glu	Glu	Glu	Glu	Val	Glu		
			850					855						860			
gag	gag	gag	gga	gaa	gag	gat	gtt	ttc	acc	gag	aaa	gcc	tca	cct	gat	2700	
Glu	Glu	Glu	Gly	Glu	Glu	Asp	Val	Phe	Thr	Glu	Lys	Ala	Ser	Pro	Asp		
			865				870					875					
atg	gat	ggg	tac	cca	gca	tta	aag	gtg	gac	aaa	gag	acc	aac	acg	gag	2748	
Met	Asp	Gly	Tyr	Pro	Ala	Leu	Lys	Val	Asp	Lys	Glu	Thr	Asn	Thr	Glu		
		880				885					890						
acc	ccg	gcc	cca	tcc	ccc	aca	gtg	gtg	cga	cct	aag	gac	cgg	aga	gtg	2796	
Thr	Pro	Ala	Pro	Ser	Pro	Thr	Val	Val	Arg	Pro	Lys	Asp	Arg	Arg	Val		
					900					905					910		

ggc acc ccg tcc cag ggg cca ttt ctt cga ggg agc acc atc atc cgc Gly Thr Pro Ser Gln Gly Pro Phe Leu Arg Gly Ser Thr Ile Ile Arg 915 920 925	2844
tct aag acc ttc tcc cca gga ccc cag agc cag tac gtg tgc cgg ctg Ser Lys Thr Phe Ser Pro Gly Pro Gln Ser Gln Tyr Val Cys Arg Leu 930 935 940	2892
aat cgg agt gat agt gac agc tcc act ctg tcc aaa aag cca cct ttt Asn Arg Ser Asp Ser Asp Ser Ser Thr Leu Ser Lys Lys Pro Pro Phe 945 950 955	2940
gtt cga aac tcc ctg gag cga cgc agc gtc cgg atg aag cgg cct tcc Val Arg Asn Ser Leu Glu Arg Arg Ser Val Arg Met Lys Arg Pro Ser 960 965 970	2988
tcg gtc aag tcg ctg cgc tcc gag cgt ctg atc cgt acc tcg ctg gac Ser Val Lys Ser Leu Arg Ser Glu Arg Leu Ile Arg Thr Ser Leu Asp 975 980 985 990	3036
ctg gag tta gac ctg cag gcg aca aga acc tgg cac agc caa ctg acc Leu Glu Leu Asp Leu Gln Ala Thr Arg Thr Trp His Ser Gln Leu Thr 995 1000 1005	3084
cag gag atc tcg gtg ctg aag gag ctc aag gag cag ctg gaa caa Gln Glu Ile Ser Val Leu Lys Glu Leu Lys Glu Gln Leu Glu Gln 1010 1015 1020	3129
gcc aag agc cac ggg gag aag gag ctg cca cag tgg ttg cgt gag Ala Lys Ser His Gly Glu Lys Glu Leu Pro Gln Trp Leu Arg Glu 1025 1030 1035	3174
gac gag cgt ttc cgc ctg ctg ctg agg atg ctg gag aag cgg cag Asp Glu Arg Phe Arg Leu Leu Leu Arg Met Leu Glu Lys Arg Gln 1040 1045 1050	3219
atg gac cga gcg gag cac aag ggt gag ctt cag aca gac aag atg Met Asp Arg Ala Glu His Lys Gly Glu Leu Gln Thr Asp Lys Met 1055 1060 1065	3264
atg agg gca gct gcc aag gat gtg cac agg ctc cga ggc cag agc Met Arg Ala Ala Ala Lys Asp Val His Arg Leu Arg Gly Gln Ser 1070 1075 1080	3309
tgt aag gaa ccc cca gaa gtt cag tct ttc agg gag aag atg gca Cys Lys Glu Pro Pro Glu Val Gln Ser Phe Arg Glu Lys Met Ala 1085 1090 1095	3354
ttt ttc acc cgg cct cgg atg aat atc cca gct ctc tct gca gat Phe Phe Thr Arg Pro Arg Met Asn Ile Pro Ala Leu Ser Ala Asp 1100 1105 1110	3399
gac gtc taatgccag aaaagtattt cctttgttcc actgaccagg ctgtgaacat Asp Val	3455
tgactgtggc taaagtattt tatgtggtgt tatatgaagg tactgagtca caagtccctt	3515
agtgctcttg ttggtttgaa gatgaaccga ctttttagtt tgggtcctac tgttgttatt	3575
aaaaacagaa caaaaacaaa acacacacac acacaaaaac agaaacaaaa aaaaccagca	3635
ttaaaataat aagattgtat agtttgtata tttaggagtg tatttttggg aaagaaaatt	3695
taaatgaact aaagcagtat tgagttgctg ctcttcttaa aatcgtttag attttttttg	3755

gtttgtacag ctccaccttt tagaggtctt actgcaataa gaagtaatgc ctgggggacg 3815  
 gtaatcctaa taggacgtcc cgcacttgtc acagtacagc taatttttcc tagttaacat 3875  
 attttgtaca atattaaaaa aatgcacaga aaccattggg ggggattcag aggtgcatcc 3935  
 acggatcttc ttgagctgtg acgtgttttt atgtggctgc ccaacgtgga gcgggcagtg 3995  
 tgataggctg ggtgggctaa gcagcctagt ctatgtgggt gacaggccac gctggtctca 4055  
 gatgccagtg gaagccacta acatgagtga ggggagggct gtggggaact ccattcagtt 4115  
 ttatctccat caataaagtg gcctttcaaa aagaaaaaaa aaaaaaaaaa aaaaaa 4171

<210> 14

<211> 1113

<212> PRT

<213> NM\_015238 KIAA0869, KIBRA

<400> 14

Met Pro Arg Pro Glu Leu Pro Leu Pro Glu Gly Trp Glu Glu Ala Arg  
 1 5 10 15

Asp Phe Asp Gly Lys Val Tyr Tyr Ile Asp His Thr Asn Arg Thr Thr  
 20 25 30

Ser Trp Ile Asp Pro Arg Asp Arg Tyr Thr Lys Pro Leu Thr Phe Ala  
 35 40 45

Asp Cys Ile Ser Asp Glu Leu Pro Leu Gly Trp Glu Glu Ala Tyr Asp  
 50 55 60

Pro Gln Val Gly Asp Tyr Phe Ile Asp His Asn Thr Lys Thr Thr Gln  
 65 70 75 80

Ile Glu Asp Pro Arg Val Gln Trp Arg Arg Glu Gln Glu His Met Leu  
 85 90 95

Lys Asp Tyr Leu Val Val Ala Gln Glu Ala Leu Ser Ala Gln Lys Glu  
 100 105 110

Ile Tyr Gln Val Lys Gln Gln Arg Leu Glu Leu Ala Gln Gln Glu Tyr  
 115 120 125

Gln Gln Leu His Ala Val Trp Glu His Lys Leu Gly Ser Gln Val Ser  
 130 135 140

Leu Val Ser Gly Ser Ser Ser Ser Lys Tyr Asp Pro Glu Ile Leu  
 145 150 155 160

Lys Ala Glu Ile Ala Thr Ala Lys Ser Arg Val Asn Lys Leu Lys Arg  
165 170 175

Glu Met Val His Leu Gln His Glu Leu Gln Phe Lys Glu Arg Gly Phe  
180 185 190

Gln Thr Leu Lys Lys Ile Asp Lys Lys Met Ser Asp Ala Gln Gly Ser  
195 200 205

Tyr Lys Leu Asp Glu Ala Gln Ala Val Leu Arg Glu Thr Lys Ala Ile  
210 215 220

Lys Lys Ala Ile Thr Cys Gly Glu Lys Glu Lys Gln Asp Leu Ile Lys  
225 230 235 240

Ser Leu Ala Met Leu Lys Asp Gly Phe Arg Thr Asp Arg Gly Ser His  
245 250 255

Ser Asp Leu Trp Ser Ser Ser Ser Ser Leu Glu Ser Ser Ser Phe Pro  
260 265 270

Leu Pro Lys Gln Tyr Leu Asp Val Ser Ser Gln Thr Asp Ile Ser Gly  
275 280 285

Ser Phe Gly Ile Asn Ser Asn Asn Gln Leu Ala Glu Lys Val Arg Leu  
290 295 300

Arg Leu Arg Tyr Glu Glu Ala Lys Arg Arg Ile Ala Asn Leu Lys Ile  
305 310 315 320

Gln Leu Ala Lys Leu Asp Ser Glu Ala Trp Pro Gly Val Leu Asp Ser  
325 330 335

Glu Arg Asp Arg Leu Ile Leu Ile Asn Glu Lys Glu Glu Leu Leu Lys  
340 345 350

Glu Met Arg Phe Ile Ser Pro Arg Lys Trp Thr Gln Gly Glu Val Glu  
355 360 365

Gln Leu Glu Met Ala Arg Lys Arg Leu Glu Lys Asp Leu Gln Ala Ala  
370 375 380

Arg Asp Thr Gln Ser Lys Ala Leu Thr Glu Arg Leu Lys Leu Asn Ser  
385 390 395 400

Lys Arg Asn Gln Leu Val Arg Glu Leu Glu Glu Ala Thr Arg Gln Val  
405 410 415

Ala Thr Leu His Ser Gln Leu Lys Ser Leu Ser Ser Ser Met Gln Ser  
420 425 430

Leu Ser Ser Gly Ser Ser Pro Gly Ser Leu Thr Ser Ser Arg Gly Ser  
435 440 445

Leu Val Ala Ser Ser Leu Asp Ser Ser Thr Ser Ala Ser Phe Thr Asp  
450 455 460

Leu Tyr Tyr Asp Pro Phe Glu Gln Leu Asp Ser Glu Leu Gln Ser Lys  
465 470 475 480

Val Glu Phe Leu Leu Leu Glu Gly Ala Thr Gly Phe Arg Pro Ser Gly  
485 490 495

Cys Ile Thr Thr Ile His Glu Asp Glu Val Ala Lys Thr Gln Lys Ala  
500 505 510

Glu Gly Gly Gly Arg Leu Gln Ala Leu Arg Ser Leu Ser Gly Thr Pro  
515 520 525

Lys Ser Met Thr Ser Leu Ser Pro Arg Ser Ser Leu Ser Ser Pro Ser  
530 535 540

Pro Pro Cys Ser Pro Leu Met Ala Asp Pro Leu Leu Ala Gly Asp Ala  
545 550 555 560

Phe Leu Asn Ser Leu Glu Phe Glu Asp Pro Glu Leu Ser Ala Thr Leu  
565 570 575

Cys Glu Leu Ser Leu Gly Asn Ser Ala Gln Glu Arg Tyr Arg Leu Glu  
580 585 590

Glu Pro Gly Thr Glu Gly Lys Gln Leu Gly Gln Ala Val Asn Thr Ala  
595 600 605

Gln Gly Cys Gly Leu Lys Val Ala Cys Val Ser Ala Ala Val Ser Asp  
610 615 620

Glu Ser Val Ala Gly Asp Ser Gly Val Tyr Glu Ala Ser Val Gln Arg  
625 630 635 640

Leu Gly Ala Ser Glu Ala Ala Ala Phe Asp Ser Asp Glu Ser Glu Ala  
645 650 655

Val Gly Ala Thr Arg Ile Gln Ile Ala Leu Lys Tyr Asp Glu Lys Asn  
660 665 670

Lys Gln Phe Ala Ile Leu Ile Ile Gln Leu Ser Asn Leu Ser Ala Leu  
675 680 685

Leu Gln Gln Gln Asp Gln Lys Val Asn Ile Arg Val Ala Val Leu Pro  
690 695 700

Cys Ser Glu Ser Thr Thr Cys Leu Phe Arg Thr Arg Pro Leu Asp Ala  
705 710 715 720

Ser Asp Thr Leu Val Phe Asn Glu Val Phe Trp Val Ser Met Ser Tyr  
725 730 735

Pro Ala Leu His Gln Lys Thr Leu Arg Val Asp Val Cys Thr Thr Asp  
740 745 750

Arg Ser His Leu Glu Glu Cys Leu Gly Gly Ala Gln Ile Ser Leu Ala  
755 760 765

Glu Val Cys Arg Ser Gly Glu Arg Ser Thr Arg Trp Tyr Asn Leu Leu  
770 775 780

Ser Tyr Lys Tyr Leu Lys Lys Gln Ser Arg Glu Leu Lys Pro Val Gly  
785 790 795 800

Val Met Ala Pro Ala Ser Gly Pro Ala Ser Thr Asp Ala Val Ser Ala  
805 810 815

Leu Leu Glu Gln Thr Ala Val Glu Leu Glu Lys Arg Gln Glu Gly Arg  
820 825 830

Ser Ser Thr Gln Thr Leu Glu Asp Ser Trp Arg Tyr Glu Glu Thr Ser  
835 840 845

Glu Asn Glu Ala Val Ala Glu Glu Glu Glu Glu Val Glu Glu Glu  
850 855 860

Glu Gly Glu Glu Asp Val Phe Thr Glu Lys Ala Ser Pro Asp Met Asp  
865 870 875 880

Gly Tyr Pro Ala Leu Lys Val Asp Lys Glu Thr Asn Thr Glu Thr Pro  
885 890 895

Ala Pro Ser Pro Thr Val Val Arg Pro Lys Asp Arg Arg Val Gly Thr  
900 905 910

Pro Ser Gln Gly Pro Phe Leu Arg Gly Ser Thr Ile Ile Arg Ser Lys  
915 920 925

Thr Phe Ser Pro Gly Pro Gln Ser Gln Tyr Val Cys Arg Leu Asn Arg  
930 935 940

Ser Asp Ser Asp Ser Ser Thr Leu Ser Lys Lys Pro Pro Phe Val Arg  
945 950 955 960

Asn Ser Leu Glu Arg Arg Ser Val Arg Met Lys Arg Pro Ser Ser Val  
965 970 975

Lys Ser Leu Arg Ser Glu Arg Leu Ile Arg Thr Ser Leu Asp Leu Glu  
 980 985 990

Leu Asp Leu Gln Ala Thr Arg Thr Trp His Ser Gln Leu Thr Gln Glu  
 995 1000 1005

Ile Ser Val Leu Lys Glu Leu Lys Glu Gln Leu Glu Gln Ala Lys  
 1010 1015 1020

Ser His Gly Glu Lys Glu Leu Pro Gln Trp Leu Arg Glu Asp Glu  
 1025 1030 1035

Arg Phe Arg Leu Leu Leu Arg Met Leu Glu Lys Arg Gln Met Asp  
 1040 1045 1050

Arg Ala Glu His Lys Gly Glu Leu Gln Thr Asp Lys Met Met Arg  
 1055 1060 1065

Ala Ala Ala Lys Asp Val His Arg Leu Arg Gly Gln Ser Cys Lys  
 1070 1075 1080

Glu Pro Pro Glu Val Gln Ser Phe Arg Glu Lys Met Ala Phe Phe  
 1085 1090 1095

Thr Arg Pro Arg Met Asn Ile Pro Ala Leu Ser Ala Asp Asp Val  
 1100 1105 1110

<210> 15

<211> 1294

<212> DNA

<213> NM\_001306 CLDN3, claudin 3

<220>

<221> CDS

<222> (198)..(857)

<223>

<400> 15

caaagccaca ggcaggtgca ggcgcagccg cggcgagagc gtatggagcc gagccgtag 60

cgcgcgccgt cggtgagtca gtccgtccgt ccgtccgtcc gtcggggcgc cgcagctccc 120

gccaggccca gcggccccgg ccctcgtct cccgcaccc ggagccaccc ggtggagcgg 180

gccttgccgc ggcagcc atg tcc atg ggc ctg gag atc acg ggc acc gcg 230  
 Met Ser Met Gly Leu Glu Ile Thr Gly Thr Ala

1	5	10	
ctg gcc gtg ctg ggc tgg ctg ggc acc atc gtg tgc tgc gcg ttg ccc Leu Ala Val Leu Gly Trp Leu Gly Thr Ile Val Cys Cys Ala Leu Pro 15 20 25			278
atg tgg cgc gtg tgc gcc ttc atc ggc agc aac atc atc acg tgc cag Met Trp Arg Val Ser Ala Phe Ile Gly Ser Asn Ile Ile Thr Ser Gln 30 35 40			326
aac atc tgg gag ggc ctg tgg atg aac tgc gtg gtg cag agc acc ggc Asn Ile Trp Glu Gly Leu Trp Met Asn Cys Val Val Gln Ser Thr Gly 45 50 55			374
cag atg cag tgc aag gtg tac gac tgc ctg ctg gca ctg cca cag gac Gln Met Gln Cys Lys Val Tyr Asp Ser Leu Leu Ala Leu Pro Gln Asp 60 65 70 75			422
ctt cag gcg gcc cgc gcc ctc atc gtg gtg gcc atc ctg ctg gcc gcc Leu Gln Ala Ala Arg Ala Leu Ile Val Val Ala Ile Leu Leu Ala Ala 80 85 90			470
ttc ggg ctg cta gtg gcg ctg gtg ggc gcc cag tgc acc aac tgc gtg Phe Gly Leu Leu Val Ala Leu Val Gly Ala Gln Cys Thr Asn Cys Val 95 100 105			518
cag gac gac acg gcc aag gcc aag atc acc atc gtg gca gcc gtg ctg Gln Asp Asp Thr Ala Lys Ala Lys Ile Thr Ile Val Ala Gly Val Leu 110 115 120			566
ttc ctt ctc gcc gcc ctg ctc acc ctc gtg ccg gtg tcc tgg tgc gcc Phe Leu Leu Ala Ala Leu Leu Thr Leu Val Pro Val Ser Trp Ser Ala 125 130 135			614
aac acc att atc cgg gac ttc tac aac ccc gtg gtg ccc gag gcg cag Asn Thr Ile Ile Arg Asp Phe Tyr Asn Pro Val Val Pro Glu Ala Gln 140 145 150 155			662
aag cgc gag atg ggc gcg gcc ctg tac gtg gcc tgg gcg gcc gcg gcg Lys Arg Glu Met Gly Ala Gly Leu Tyr Val Gly Trp Ala Ala Ala Ala 160 165 170			710
ctg cag ctg ctg ggg gcc gcg ctg ctc tgc tgc tgc tgt ccc cca cgc Leu Gln Leu Leu Gly Gly Ala Leu Leu Cys Cys Ser Cys Pro Pro Arg 175 180 185			758
gag aag aag tac acg gcc acc aag gtc gtc tac tcc gcg ccg cgc tcc Glu Lys Lys Tyr Thr Ala Thr Lys Val Val Tyr Ser Ala Pro Arg Ser 190 195 200			806
acc ggc ccg gga gcc agc ctg ggc aca ggc tac gac cgc aag gac tac Thr Gly Pro Gly Ala Ser Leu Gly Thr Gly Tyr Asp Arg Lys Asp Tyr 205 210 215			854
gtc taagggacag acgcagggag accccaccac caccaccac accaacacca Val 220			907
ccaccaccac cgcgagctgg agcgcgcacc aggccatcca gcgtgcagcc ttgcctcgga			967
ggccagccca cccccagaag ccaggaagcc cccgcgctgg actggggcag cttccccagc			1027
agccacggct ttgcgggccc gccagtcgac ttcggggccc agggaccaac ctgcatggac			1087
tgtgaaacct cacccttctg gagcacgggg cctgggtgac cgccaatact tgaccacccc			1147
gtcgagcccc atcgggccgc tgcccccatt ctcgcgctgg gcagggaccg gcagccctgg			1207



aaggggcact tgatattttt caataaaagc ctttcgtttt gcaaaaaaaa aaaaaaaaaa 1267

aaaaaaaaaa aaaaaaaaaa aaaaaaa 1294

<210> 16

<211> 220

<212> PRT

<213> NM\_001306 CLDN3, claudin 3

<400> 16

Met Ser Met Gly Leu Glu Ile Thr Gly Thr Ala Leu Ala Val Leu Gly  
1 5 10 15

Trp Leu Gly Thr Ile Val Cys Cys Ala Leu Pro Met Trp Arg Val Ser  
20 25 30

Ala Phe Ile Gly Ser Asn Ile Ile Thr Ser Gln Asn Ile Trp Glu Gly  
35 40 45

Leu Trp Met Asn Cys Val Val Gln Ser Thr Gly Gln Met Gln Cys Lys  
50 55 60

Val Tyr Asp Ser Leu Leu Ala Leu Pro Gln Asp Leu Gln Ala Ala Arg  
65 70 75 80

Ala Leu Ile Val Val Ala Ile Leu Leu Ala Ala Phe Gly Leu Leu Val  
85 90 95

Ala Leu Val Gly Ala Gln Cys Thr Asn Cys Val Gln Asp Asp Thr Ala  
100 105 110

Lys Ala Lys Ile Thr Ile Val Ala Gly Val Leu Phe Leu Leu Ala Ala  
115 120 125

Leu Leu Thr Leu Val Pro Val Ser Trp Ser Ala Asn Thr Ile Ile Arg  
130 135 140

Asp Phe Tyr Asn Pro Val Val Pro Glu Ala Gln Lys Arg Glu Met Gly  
145 150 155 160

Ala Gly Leu Tyr Val Gly Trp Ala Ala Ala Ala Leu Gln Leu Leu Gly  
165 170 175

Gly Ala Leu Leu Cys Cys Ser Cys Pro Pro Arg Glu Lys Lys Tyr Thr  
180 185 190

Ala Thr Lys Val Val Tyr Ser Ala Pro Arg Ser Thr Gly Pro Gly Ala  
195 200 205

Ser Leu Gly Thr Gly Tyr Asp Arg Lys Asp Tyr Val  
 210 215 220

<210> 17

<211> 1853

<212> DNA

<213> NM\_022454 SOX17

<220>

<221> CDS

<222> (205) .. (1446)

<223>

<400> 17

gcagtgtcac taggccggct gggggccctg ggtacgctgt agaccagacc gcgacaggcc 60

agaacacggg cgccggcttc gggccgggag acccgcgag ccctcggggc atctcagtgc 120

ctcactcccc accccctccc ccgggtcggg ggaggcggcg cgtccggcgg agggttgagg 180

ggagcggggc aggcctggag cgcc atg agc agc ccg gat gcg gga tac gcc 231  
 Met Ser Ser Pro Asp Ala Gly Tyr Ala  
 1 5

agt gac gac cag agc cag acc cag agc gcg ctg ccc gcg gtg atg gcc 279  
 Ser Asp Asp Gln Ser Gln Thr Gln Ser Ala Leu Pro Ala Val Met Ala  
 10 15 20 25

ggg ctg ggc ccc tgc ccc tgg gcc gag tgc ctg agc ccc atc ggg gac 327  
 Gly Leu Gly Pro Cys Pro Trp Ala Glu Ser Leu Ser Pro Ile Gly Asp  
 30 35 40

atg aag gtg aag ggc gag gcg ccg gcg aac agc gga gca ccg gcc ggg 375  
 Met Lys Val Lys Gly Glu Ala Pro Ala Asn Ser Gly Ala Pro Ala Gly  
 45 50 55

gcc gcg ggc cga gcc aag ggc gag tcc cgt atc cgg cgg ccg atg aac 423  
 Ala Ala Gly Arg Ala Lys Gly Glu Ser Arg Ile Arg Arg Pro Met Asn  
 60 65 70

gct ttc atg gtg tgg gct aag gac gag cgc aag cgg ctg gcg cag cag 471

Ala Phe Met Val Trp Ala Lys Asp Glu Arg Lys Arg Leu Ala Gln Gln  
 75 80 85

aat cca gac ctg cac aac gcc gag ttg agc aag atg ctg ggc aag tgc 519  
 Asn Pro Asp Leu His Asn Ala Glu Leu Ser Lys Met Leu Gly Lys Ser  
 90 95 100 105

tgg aag gcg ctg acg ctg gcg gag aag cgg ccc ttc gtg gag gag gca 567  
 Trp Lys Ala Leu Thr Leu Ala Glu Lys Arg Pro Phe Val Glu Glu Ala  
 110 115 120

gag cgg ctg cgc gtg cag cac atg cag gac cac ccc aac tac aag tac 615  
 Glu Arg Leu Arg Val Gln His Met Gln Asp His Pro Asn Tyr Lys Tyr  
 125 130 135

cgg cgc cgg cgg cgc aag cag gtg aag cgg ctg aag cgg gtg gag ggc Arg Pro Arg Arg Arg Lys Gln Val Lys Arg Leu Lys Arg Val Glu Gly 140 145 150	663
ggc ttc ctg cac ggc ctg gct gag cgc cag gcg gcc gcg ctg ggc ccc Gly Phe Leu His Gly Leu Ala Glu Pro Gln Ala Ala Ala Leu Gly Pro 155 160 165	711
gag ggc ggc cgc gtg gcc atg gac ggc ctg ggc ctc cag ttc ccc gag Glu Gly Gly Arg Val Ala Met Asp Gly Leu Gly Leu Gln Phe Pro Glu 170 175 180 185	759
cag ggc ttc ccc gcc ggc ccg ccg ctg ctg cct ccg cac atg ggc ggc Gln Gly Phe Pro Ala Gly Pro Pro Leu Leu Pro Pro His Met Gly Gly 190 195 200	807
cac tac cgc gac tgc cag agt ctg ggc gcg cct ccg ctc gac ggc tac His Tyr Arg Asp Cys Gln Ser Leu Gly Ala Pro Pro Leu Asp Gly Tyr 205 210 215	855
ccg ttg ccc acg ccc gac acg tcc ccg ctg gac ggc gtg gac ccc gac Pro Leu Pro Thr Pro Asp Thr Ser Pro Leu Asp Gly Val Asp Pro Asp 220 225 230	903
ccg gct ttc ttc gcc gcc ccg atg ccc ggc gac tgc ccg gcg gcc ggc Pro Ala Phe Phe Ala Ala Pro Met Pro Gly Asp Cys Pro Ala Ala Gly 235 240 245	951
acc tac agc tac gcg cag gtc tcg gac tac gct ggc ccc ccg gag cct Thr Tyr Ser Tyr Ala Gln Val Ser Asp Tyr Ala Gly Pro Pro Glu Pro 250 255 260 265	999
ccc gcc ggt ccc atg cac ccc cga ctc ggc cca gag ccc gcg ggt ccc Pro Ala Gly Pro Met His Pro Arg Leu Gly Pro Glu Pro Ala Gly Pro 270 275 280	1047
tcg att ccg ggc ctc ctg gcg cca ccc agc gcc ctt cac gtg tac tac Ser Ile Pro Gly Leu Leu Ala Pro Pro Ser Ala Leu His Val Tyr Tyr 285 290 295	1095
ggc gcg atg ggc tcg ccc ggc gcg ggc ggc ggc gcg ggc ttc cag atg Gly Ala Met Gly Ser Pro Gly Ala Gly Gly Gly Arg Gly Phe Gln Met 300 305 310	1143
cag ccg caa cac cag cac cag cac cag cac cag ccc ccg ggc Gln Pro Gln His Gln His Gln His Gln His Gln His Pro Pro Gly 315 320 325	1191
ccc gga cag ccg tcg ccc cct ccg gag gca ctg ccc tgc ccg gac ggc Pro Gly Gln Pro Ser Pro Pro Pro Glu Ala Leu Pro Cys Arg Asp Gly 330 335 340 345	1239
acg gac ccc agt cag ccc gcc gag ctc ctc ggc gag gtg gac cgc acg Thr Asp Pro Ser Gln Pro Ala Glu Leu Leu Gly Glu Val Asp Arg Thr 350 355 360	1287
gaa ttt gaa cag tat ctg cac ttc gtg tgc aag cct gag atg ggc ctc Glu Phe Glu Gln Tyr Leu His Phe Val Cys Lys Pro Glu Met Gly Leu 365 370 375	1335
ccc tac cag ggc cat gac tcc ggt gtg aat ctc ccc gac agc cac ggc Pro Tyr Gln Gly His Asp Ser Gly Val Asn Leu Pro Asp Ser His Gly 380 385 390	1383
gcc att tcc tcg gtg gtg tcc gac gcc agc tcc gcg gta tat tac tgc	1431

Ala Ile Ser Ser Val Val Ser Asp Ala Ser Ser Ala Val Tyr Tyr Cys  
395 400 405

aac tat cct gac gtg tgacaggtcc ctgatccgcc ccagcctgca ggccagaagc 1486  
Asn Tyr Pro Asp Val  
410

agtgttacac acttctctgga ggagctaagg aaatcctcag actcctgggt ttttgttgtt 1546

gctgttgttg ttttttaaaa ggtgtgttg catataattt atggttaattt atttgtctg 1606

ccacttgaac agtttggggg ggtgaggttt catttaaaat ttgttcagag atttgtttcc 1666

catagtgtgga ttgtcaaaac cctatttcca agttcaagtt aactagcttt gaatgtgtcc 1726

caaaacagct tcttccattt cctgaaagt ttattgatcaa agaaatgttg tcttgggtgt 1786

gttttttcaa tcttctaaaa aataaaatct ggaatcctga aaaa'aaaaa aaaaaaaaaa 1846

aaaaaaaa 1853

<210> 18

<211> 414

<212> PRT

<213> NM\_022454 SOX17

<400> 18

Met Ser Ser Pro Asp Ala Gly Tyr Ala Ser Asp Asp Gln Ser Gln Thr  
1 5 10 15

Gln Ser Ala Leu Pro Ala Val Met Ala Gly Leu Gly Pro Cys Pro Trp  
20 25 30

Ala Glu Ser Leu Ser Pro Ile Gly Asp Met Lys Val Lys Gly Glu Ala  
35 40 45

Pro Ala Asn Ser Gly Ala Pro Ala Gly Ala Ala Gly Arg Ala Lys Gly  
50 55 60

Glu Ser Arg Ile Arg Arg Pro Met Asn Ala Phe Met Val Trp Ala Lys  
65 70 75 80

Asp Glu Arg Lys Arg Leu Ala Gln Gln Asn Pro Asp Leu His Asn Ala  
85 90 95

Glu Leu Ser Lys Met Leu Gly Lys Ser Trp Lys Ala Leu Thr Leu Ala  
100 105 110

Glu Lys Arg Pro Phe Val Glu Glu Ala Glu Arg Leu Arg Val Gln His  
115 120 125

Met Gln Asp His Pro Asn Tyr Lys Tyr Arg Pro Arg Arg Arg Lys Gln  
130 135 140

Val Lys Arg Leu Lys Arg Val Glu Gly Gly Phe Leu His Gly Leu Ala  
145 150 155 160

Glu Pro Gln Ala Ala Ala Leu Gly Pro Glu Gly Gly Arg Val Ala Met  
165 170 175

Asp Gly Leu Gly Leu Gln Phe Pro Glu Gln Gly Phe Pro Ala Gly Pro  
180 185 190

Pro Leu Leu Pro Pro His Met Gly Gly His Tyr Arg Asp Cys Gln Ser  
195 200 205

Leu Gly Ala Pro Pro Leu Asp Gly Tyr Pro Leu Pro Thr Pro Asp Thr  
210 215 220

Ser Pro Leu Asp Gly Val Asp Pro Asp Pro Ala Phe Phe Ala Ala Pro  
225 230 235 240

Met Pro Gly Asp Cys Pro Ala Ala Gly Thr Tyr Ser Tyr Ala Gln Val  
245 250 255

Ser Asp Tyr Ala Gly Pro Pro Glu Pro Pro Ala Gly Pro Met His Pro  
260 265 270

Arg Leu Gly Pro Glu Pro Ala Gly Pro Ser Ile Pro Gly Leu Leu Ala  
275 280 285

Pro Pro Ser Ala Leu His Val Tyr Tyr Gly Ala Met Gly Ser Pro Gly  
290 295 300

Ala Gly Gly Gly Arg Gly Phe Gln Met Gln Pro Gln His Gln His Gln  
305 310 315 320

His Gln His Gln His His Pro Pro Gly Pro Gly Gln Pro Ser Pro Pro  
325 330 335

Pro Glu Ala Leu Pro Cys Arg Asp Gly Thr Asp Pro Ser Gln Pro Ala  
340 345 350

Glu Leu Leu Gly Glu Val Asp Arg Thr Glu Phe Glu Gln Tyr Leu His  
355 360 365

Phe Val Cys Lys Pro Glu Met Gly Leu Pro Tyr Gln Gly His Asp Ser  
370 375 380

Gly Val Asn Leu Pro Asp Ser His Gly Ala Ile Ser Ser Val Val Ser  
385 390 395 400

Asp Ala Ser Ser Ala Val Tyr Tyr Cys Asn Tyr Pro Asp Val  
405 410

&lt;210&gt; 19

&lt;211&gt; 3702

&lt;212&gt; DNA

&lt;213&gt; NM\_005682 GPR56

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (159)..(2237)

&lt;223&gt;

&lt;400&gt; 19

tggagggtct	cgctctgtca	cacaggctgg	agtgcagtgg	tgtgatcttg	gctcatcgta	60
------------	------------	------------	------------	------------	------------	----

acctccacct	ccccgggttca	agtgattctc	atgcctcagc	ctcccagta	gctgggatta	120
------------	-------------	------------	------------	-----------	------------	-----

caggtggtga	cttccaagag	tgactccgtc	ggaggaaa	atg act ccc	cag tgc ctg	176
				Met Thr Pro	Gln Ser Leu	
				1	5	

ctg cag acg	aca ctg ttc	ctg ctg agt	ctg ctc ttc	ctg gtc	caa ggt	224
Leu Gln Thr	Thr Leu Phe	Leu Leu Ser	Leu Leu Phe	Leu Val	Gln Gly	
	10		15		20	

gcc cac ggc	agg ggc cac	agg gaa gac	ttt cgc ttc	tgc agc	cag cgg	272
Ala His Gly	Arg Gly His	Arg Glu Asp	Phe Arg Phe	Cys Ser	Gln Arg	
	25		30		35	

aac cag aca	cac agg agc	agc ctc cac	tac aaa ccc	aca cca	gac ctg	320
Asn Gln Thr	His Arg Ser	Ser Leu His	Tyr Lys Pro	Thr Pro	Asp Leu	
	40		45		50	

cgc atc tcc	atc gag aac	tcc gaa gag	gcc ctc aca	gtc cat	gcc cct	368
Arg Ile Ser	Ile Glu Asn	Ser Glu Glu	Ala Leu Thr	Val His	Ala Pro	
	55		60		65	

ttc cct gca	gcc cac cct	gct tcc cga	tcc ttc cct	gac ccc	agg ggc	416
Phe Pro Ala	Ala His Pro	Ala Ser Arg	Ser Phe Pro	Asp Pro	Arg Gly	
	75		80		85	

ctc tac cac	ttc tgc ctc	tac tgg aac	cga cat gct	ggg aga	tta cat	464
Leu Tyr His	Phe Cys Leu	Tyr Trp Asn	Arg His Ala	Gly Arg	Leu His	
	90		95		100	

ctt ctc tat	ggc aag cgt	gac ttc ttg	ctg agt gac	aaa gcc	tct agc	512
Leu Leu Tyr	Gly Lys Arg	Asp Phe Leu	Leu Ser Asp	Lys Ala	Ser Ser	
	105		110		115	

ctc ctc tgc	ttc cag cac	cag gag gag	agc ctg gct	cag ggc	ccc ccg	560
Leu Leu Cys	Phe Gln His	Gln Glu Glu	Ser Leu Ala	Gln Gly	Pro Pro	
	120		125		130	

ctg tta gcc	act tct gtc	acc tcc tgg	tgg agc cct	cag aac	atc agc	608
Leu Leu Ala	Thr Ser Val	Thr Ser Trp	Trp Ser Pro	Gln Asn	Ile Ser	
	135		140		145	

ctg ccc agt	gcc gcc agc	ttc acc ttc	tcc ttc cac	agt cct	ccc cac	656
-------------	-------------	-------------	-------------	---------	---------	-----

Leu Pro Ser Ala Ala Ser Phe Thr Phe Ser Phe His Ser Pro Pro His	
155 160 165	
acg gcc gct cac aat gcc tcg gtg gac atg tgc gag ctc aaa agg gac	704
Thr Ala Ala His Asn Ala Ser Val Asp Met Cys Glu Leu Lys Arg Asp	
170 175 180	
ctc cag ctg ctc agc cag ttc ctg aag cat ccc cag aag gcc tca agg	752
Leu Gln Leu Leu Ser Gln Phe Leu Lys His Pro Gln Lys Ala Ser Arg	
185 190 195	
agg ccc tcg gct gcc ccc gcc agc cag cag ttg cag agc ctg gag tcg	800
Arg Pro Ser Ala Ala Pro Ala Ser Gln Gln Leu Gln Ser Leu Glu Ser	
200 205 210	
aaa ctg acc tct gtg aga ttc atg ggg gac atg gtg tcc ttc gag gag	848
Lys Leu Thr Ser Val Arg Phe Met Gly Asp Met Val Ser Phe Glu Glu	
215 220 225 230	
gac cgg atc aac gcc acg gtg tgg aag ctc cag ccc aca gcc ggc ctc	896
Asp Arg Ile Asn Ala Thr Val Trp Lys Leu Gln Pro Thr Ala Gly Leu	
235 240 245	
cag gac ctg cac atc cac tcc cgg cag gag gag gag cag agc gag atc	944
Gln Asp Leu His Ile His Ser Arg Gln Glu Glu Gln Ser Glu Ile	
250 255 260	
atg gag tac tcg gtg ctg ctg cct cga aca ctc ttc cag agg acg aaa	992
Met Glu Tyr Ser Val Leu Leu Pro Arg Thr Leu Phe Gln Arg Thr Lys	
265 270 275	
ggc cgg agc ggg gag gct gag aag aga ctc ctc ctg gtg gac ttc agc	1040
Gly Arg Ser Gly Glu Ala Glu Lys Arg Leu Leu Leu Val Asp Phe Ser	
280 285 290	
agc caa gcc ctg ttc cag gac aag aat tcc agc cac gtc ctg ggt gag	1088
Ser Gln Ala Leu Phe Gln Asp Lys Asn Ser Ser His Val Leu Gly Glu	
295 300 305 310	
aag gtc ttg ggg att gtg gta cag aac acc aaa gta gcc aac ctc acg	1136
Lys Val Leu Gly Ile Val Val Gln Asn Thr Lys Val Ala Asn Leu Thr	
315 320 325	
gag ccc gtg gtg ctc acc ttc cag cac cag cta cag ccg aag aat gtg	1184
Glu Pro Val Val Leu Thr Phe Gln His Gln Leu Gln Pro Lys Asn Val	
330 335 340	
act ctg caa tgt gtg ttc tgg gtt gaa gac ccc aca ttg agc agc ccg	1232
Thr Leu Gln Cys Val Phe Trp Val Glu Asp Pro Thr Leu Ser Ser Pro	
345 350 355	
ggg cat tgg agc agt gct ggg tgt gag acc gtc agg aga gaa acc caa	1280
Gly His Trp Ser Ser Ala Gly Cys Glu Thr Val Arg Arg Glu Thr Gln	
360 365 370	
aca tcc tgc ttc tgc aac cac ttg acc tac ttt gca gtg ctg atg gtc	1328
Thr Ser Cys Phe Cys Asn His Leu Thr Tyr Phe Ala Val Leu Met Val	
375 380 385 390	
tcc tcg gtg gag gtg gac gcc gtg cac aag cac tac ctg agc ctc ctc	1376
Ser Ser Val Glu Val Asp Ala Val His Lys His Tyr Leu Ser Leu Leu	
395 400 405	
tcc tac gtg ggc tgt gtc gtc tct gcc ctg gcc tgc ctt gtc acc att	1424
Ser Tyr Val Gly Cys Val Val Ser Ala Leu Ala Cys Leu Val Thr Ile	
410 415 420	
gcc gcc tac ctc tgc tcc agg gtg ccc ctg ccg tgc agg agg aaa cct	1472

Ala Ala Tyr Leu Cys Ser Arg Val Pro Leu Pro Cys Arg Arg Lys Pro	
425 430 435	
cgg gac tac acc atc aag gtg cac atg aac ctg ctg ctg gcc gtc ttc	1520
Arg Asp Tyr Thr Ile Lys Val His Met Asn Leu Leu Leu Ala Val Phe	
440 445 450	
ctg ctg gac acg agc ttc ctg ctc agc gag ccg gtg gcc ctg aca ggc	1568
Leu Leu Asp Thr Ser Phe Leu Leu Ser Glu Pro Val Ala Leu Thr Gly	
455 460 465 470	
tct gag gct ggc tgc cga gcc agt gcc atc ttc ctg cac ttc tcc ctg	1616
Ser Glu Ala Gly Cys Arg Ala Ser Ala Ile Phe Leu His Phe Ser Leu	
475 480 485	
ctc acc tgc ctt tcc tgg atg ggc ctc gag ggg tac aac ctc tac cga	1664
Leu Thr Cys Leu Ser Trp Met Gly Leu Glu Gly Tyr Asn Leu Tyr Arg	
490 495 500	
ctc gtg gtg gag gtc ttt ggc acc tat gtc cct ggc tac cta ctc aag	1712
Leu Val Val Glu Val Phe Gly Thr Tyr Val Pro Gly Tyr Leu Leu Lys	
505 510 515	
ctg agc gcc atg ggc tgg ggc ttc ccc atc ttt ctg gtg acg ctg gtg	1760
Leu Ser Ala Met Gly Trp Gly Phe Pro Ile Phe Leu Val Thr Leu Val	
520 525 530	
gcc ctg gtg gat gtg gac aac tat ggc ccc atc atc ttg gct gtg cat	1808
Ala Leu Val Asp Val Asp Asn Tyr Gly Pro Ile Ile Leu Ala Val His	
535 540 545 550	
agg act cca gag ggc gtc atc tac cct tcc atg tgc tgg atc cgg gac	1856
Arg Thr Pro Glu Gly Val Ile Tyr Pro Ser Met Cys Trp Ile Arg Asp	
555 560 565	
tcc ctg gtc agc tac atc acc aac ctg ggc ctc ttc agc ctg gtg ttt	1904
Ser Leu Val Ser Tyr Ile Thr Asn Leu Gly Leu Phe Ser Leu Val Phe	
570 575 580	
ctg ttc aac atg gcc atg cta gcc acc atg gtg gtg cag atc ctg cgg	1952
Leu Phe Asn Met Ala Met Leu Ala Thr Met Val Val Gln Ile Leu Arg	
585 590 595	
ctg cgc ccc cac acc caa aag tgg tca cat gtg ctg aca ctg ctg ggc	2000
Leu Arg Pro His Thr Gln Lys Trp Ser His Val Leu Thr Leu Leu Gly	
600 605 610	
ctc agc ctg gtc ctt ggc ctg ccc tgg gcc ttg atc ttc ttc tcc ttt	2048
Leu Ser Leu Val Leu Gly Leu Pro Trp Ala Leu Ile Phe Phe Ser Phe	
615 620 625 630	
gct tct ggc acc ttc cag ctt gtc gtc ctc tac ctt ttc agc atc atc	2096
Ala Ser Gly Thr Phe Gln Leu Val Val Leu Tyr Leu Phe Ser Ile Ile	
635 640 645	
acc tcc ttc caa ggc ttc ctc atc ttc atc tgg tac tgg tcc atg cgg	2144
Thr Ser Phe Gln Gly Phe Leu Ile Phe Ile Trp Tyr Trp Ser Met Arg	
650 655 660	
ctg cag gcc cgg ggt ggc ccc tcc cct ctg aag agc aac tca gac agc	2192
Leu Gln Ala Arg Gly Gly Pro Ser Pro Leu Lys Ser Asn Ser Asp Ser	
665 670 675	
gcc agg ctc ccc atc agc tcg ggc agc acc tcg tcc agc cgc atc	2237
Ala Arg Leu Pro Ile Ser Ser Gly Ser Thr Ser Ser Arg Ile	
680 685 690	
taggcctcca gccacctgc ccatgtgatg aagcagagat gcggcctcgt cgcacactgc	2297



ctgtggcccc cgagccaggc ccagccccag gccagtcagc cgcagacttt ggaaagccca 2357  
 acgaccatgg agagatgggc cgttgccatg gtggacggac tcccgggctg ggcttttgaa 2417  
 ttggccttgg ggactactcg gctctcactc agctcccacg ggactcagaa gtgcgcgcgc 2477  
 atgctgccta ggggtactgtc cccacatctg tcccaacca gctggaggcc tggctctctc 2537  
 ttacaaaccc tgggcccagc cctcattgct gggggccagg ccttggtatc tgagggtctg 2597  
 gcaatcctt aatcctgtgc cctgcctgg gacagaaatg tggctccagt tgctctgtct 2657  
 ctgctggtca ccctgagggc actctgcac cctctgtcatt ttaacctcag gtggcaccca 2717  
 gggcgaatgg ggcccagggc agaccttcag ggccagagcc ctggcgagg agaggccctt 2777  
 tgccaggagc acagcagcag ctgcctacc tctgagccca ggccccctcc ctccctcagc 2837  
 cccccagtc tccctccatc ttccctgggg ttctctctct ctcccagggc ctccctgtct 2897  
 cttcgttcac agctgggggt ccccgattcc aatgctgttt ttggggagt ggtttccagg 2957  
 agctgcctgg tgtctgtgt aaatgtttgt ctactgcaca agcctcggcc tgcccctgag 3017  
 ccaggctcgg taccgatgcg tgggctgggc taggtccctc tgtccatctg ggcctttgta 3077  
 tgagctgcat tgcccttgtc caccctgacc aagcacacgc ctccagaggg ccctcagcct 3137  
 ctctgaagc cctcttgtgg caagaactgt ggaccatgcc agtcccgctc ggtttccatc 3197  
 ccaccactcc aaggactgag actgacctcc tctggtgaca ctggcctaga gcctgacact 3257  
 ctctaagag gttctctcca agccccaaa tagctccagg cgcctcggc cgcccatcat 3317  
 ggttaattct gtccaacaaa cacacacggg tagattgctg gcctgttgta ggtggtagg 3377  
 acacagatga ccgacctggt cactcctcct gccaacattc agtctggtat gtgaggcgtg 3437  
 cgtgaagcaa gaactcctgg agctacaggg acaggagacc atcattcctg cctgggaatc 3497  
 ctggaagact tctgcagga gtcagcgtc aatcttgacc ttgaagatgg gaaggatgtt 3557  
 ctttttacgt accaattctt ttgtcttttg atattaaaaa gaagtacatg ttcattgtag 3617  
 agaatttga aactgtagaa gagaatcaag aagaaaaata aaaatcagct gttgtaatcg 3677  
 cctagcaaaa aaaaaaaaaa aaaaa 3702

<210> 20

<211> 693

<212> PRT

<213> NM\_005682 GPR56

<400> 20

Met Thr Pro Gln Ser Leu Leu Gln Thr Thr Leu Phe Leu Leu Ser Leu  
 1 5 10 15

Leu Phe Leu Val Gln Gly Ala His Gly Arg Gly His Arg Glu Asp Phe

20 25 30  
 Arg Phe Cys Ser Gln Arg Asn Gln Thr His Arg Ser Ser Leu His Tyr  
 35 40 45  
 Lys Pro Thr Pro Asp Leu Arg Ile Ser Ile Glu Asn Ser Glu Glu Ala  
 50 55 60  
 Leu Thr Val His Ala Pro Phe Pro Ala Ala His Pro Ala Ser Arg Ser  
 65 70 75 80  
 Phe Pro Asp Pro Arg Gly Leu Tyr His Phe Cys Leu Tyr Trp Asn Arg  
 85 90 95  
 His Ala Gly Arg Leu His Leu Leu Tyr Gly Lys Arg Asp Phe Leu Leu  
 100 105 110  
 Ser Asp Lys Ala Ser Ser Leu Leu Cys Phe Gln His Gln Glu Glu Ser  
 115 120 125  
 Leu Ala Gln Gly Pro Pro Leu Leu Ala Thr Ser Val Thr Ser Trp Trp  
 130 135 140  
 Ser Pro Gln Asn Ile Ser Leu Pro Ser Ala Ala Ser Phe Thr Phe Ser  
 145 150 155 160  
 Phe His Ser Pro Pro His Thr Ala Ala His Asn Ala Ser Val Asp Met  
 165 170 175  
 Cys Glu Leu Lys Arg Asp Leu Gln Leu Leu Ser Gln Phe Leu Lys His  
 180 185 190  
 Pro Gln Lys Ala Ser Arg Arg Pro Ser Ala Ala Pro Ala Ser Gln Gln  
 195 200 205  
 Leu Gln Ser Leu Glu Ser Lys Leu Thr Ser Val Arg Phe Met Gly Asp  
 210 215 220  
 Met Val Ser Phe Glu Glu Asp Arg Ile Asn Ala Thr Val Trp Lys Leu  
 225 230 235 240  
 Gln Pro Thr Ala Gly Leu Gln Asp Leu His Ile His Ser Arg Gln Glu  
 245 250 255  
 Glu Glu Gln Ser Glu Ile Met Glu Tyr Ser Val Leu Leu Pro Arg Thr  
 260 265 270  
 Leu Phe Gln Arg Thr Lys Gly Arg Ser Gly Glu Ala Glu Lys Arg Leu  
 275 280 285  
 Leu Leu Val Asp Phe Ser Ser Gln Ala Leu Phe Gln Asp Lys Asn Ser

290 295 300

Ser His Val Leu Gly Glu Lys Val Leu Gly Ile Val Val Gln Asn Thr  
305 310 315 320

Lys Val Ala Asn Leu Thr Glu Pro Val Val Leu Thr Phe Gln His Gln  
325 330 335

Leu Gln Pro Lys Asn Val Thr Leu Gln Cys Val Phe Trp Val Glu Asp  
340 345 350

Pro Thr Leu Ser Ser Pro Gly His Trp Ser Ser Ala Gly Cys Glu Thr  
355 360 365

Val Arg Arg Glu Thr Gln Thr Ser Cys Phe Cys Asn His Leu Thr Tyr  
370 375 380

Phe Ala Val Leu Met Val Ser Ser Val Glu Val Asp Ala Val His Lys  
385 390 395 400

His Tyr Leu Ser Leu Leu Ser Tyr Val Gly Cys Val Val Ser Ala Leu  
405 410 415

Ala Cys Leu Val Thr Ile Ala Ala Tyr Leu Cys Ser Arg Val Pro Leu  
420 425 430

Pro Cys Arg Arg Lys Pro Arg Asp Tyr Thr Ile Lys Val His Met Asn  
435 440 445

Leu Leu Leu Ala Val Phe Leu Leu Asp Thr Ser Phe Leu Leu Ser Glu  
450 455 460

Pro Val Ala Leu Thr Gly Ser Glu Ala Gly Cys Arg Ala Ser Ala Ile  
465 470 475 480

Phe Leu His Phe Ser Leu Leu Thr Cys Leu Ser Trp Met Gly Leu Glu  
485 490 495

Gly Tyr Asn Leu Tyr Arg Leu Val Val Glu Val Phe Gly Thr Tyr Val  
500 505 510

Pro Gly Tyr Leu Leu Lys Leu Ser Ala Met Gly Trp Gly Phe Pro Ile  
515 520 525

Phe Leu Val Thr Leu Val Ala Leu Val Asp Val Asp Asn Tyr Gly Pro  
530 535 540

Ile Ile Leu Ala Val His Arg Thr Pro Glu Gly Val Ile Tyr Pro Ser  
545 550 555 560

Met Cys Trp Ile Arg Asp Ser Leu Val Ser Tyr Ile Thr Asn Leu Gly  
565 570 575

Leu Phe Ser Leu Val Phe Leu Phe Asn Met Ala Met Leu Ala Thr Met  
580 585 590

Val Val Gln Ile Leu Arg Leu Arg Pro His Thr Gln Lys Trp Ser His  
595 600 605

Val Leu Thr Leu Leu Gly Leu Ser Leu Val Leu Gly Leu Pro Trp Ala  
610 615 620

Leu Ile Phe Phe Ser Phe Ala Ser Gly Thr Phe Gln Leu Val Val Leu  
625 630 635 640

Tyr Leu Phe Ser Ile Ile Thr Ser Phe Gln Gly Phe Leu Ile Phe Ile  
645 650 655

Trp Tyr Trp Ser Met Arg Leu Gln Ala Arg Gly Gly Pro Ser Pro Leu  
660 665 670

Lys Ser Asn Ser Asp Ser Ala Arg Leu Pro Ile Ser Ser Gly Ser Thr  
675 680 685

Ser Ser Ser Arg Ile  
690

<210> 21

<211> 1332

<212> DNA

<213> NM\_001307 claudin 7, CLDN7

<220>

<221> CDS

<222> (427)..(1059)

<223>

<400> 21

ctgctggctc acctccgagc cacctctgct gcgcaccgca gcttcggacc tacagcccag	60
gatactttgg gacttgccgg cgctcagaaa cgcgcccaga cggcccctcc accttttgtt	120
tgcctagggt cgccgagagc gcccgagggg aaccgcctgg ccttcgggga ccaccaattt	180
tgtctggaac caccctcccg gcgtatccta ctccctgtgc cgcgaggcca tcgcttcaact	240
ggaggggtcg attttgtgtg agtttggtga caagatttgc attcacctgg cccaaacctt	300
ttttgtctct ttgggtgacc ggaaaactcc acctcaagtt ttcttttgtg gggctgcccc	360

ccaagtgtcg tttgttttac tgtaggtct cccgcccgcc gccccagtg tttctgagg	420
gcggaa atg gcc aat tcg ggc ctg cag ttg ctg ggc ttc tcc atg gcc	468
Met Ala Asn Ser Gly Leu Gln Leu Leu Gly Phe Ser Met Ala	
1 5 10	
ctg ctg ggc tgg gtg ggt ctg gtg gcc tgc acc gcc atc ccg cag tgg	516
Leu Leu Gly Trp Val Gly Leu Val Ala Cys Thr Ala Ile Pro Gln Trp	
15 20 25 30	
cag atg agc tcc tat gcg ggt gac aac atc atc acg gcc cag gcc atg	564
Gln Met Ser Ser Tyr Ala Gly Asp Asn Ile Ile Thr Ala Gln Ala Met	
35 40 45	
tac aag ggg ctg tgg atg gac tgc gtc acg cag agc acg ggg atg atg	612
Tyr Lys Gly Leu Trp Met Asp Cys Val Thr Gln Ser Thr Gly Met Met	
50 55 60	
agc tgc aaa atg tac gac tcg gtg ctc gcc ctg tcc gcg gcc ttg cag	660
Ser Cys Lys Met Tyr Asp Ser Val Leu Ala Leu Ser Ala Ala Leu Gln	
65 70 75	
gcc act cga gcc cta atg gtg gtc tcc ctg gtg ctg ggc ttc ctg gcc	708
Ala Thr Arg Ala Leu Met Val Val Ser Leu Val Leu Gly Phe Leu Ala	
80 85 90	
atg ttt gtg gcc acg atg ggc atg aag tgc acg cgc tgt ggg gga gac	756
Met Phe Val Ala Thr Met Gly Met Lys Cys Thr Arg Cys Gly Gly Asp	
95 100 105 110	
gac aaa gtg aag aag gcc cgt ata gcc atg ggt gga ggc ata att ttc	804
Asp Lys Val Lys Lys Ala Arg Ile Ala Met Gly Gly Gly Ile Ile Phe	
115 120 125	
atc gtg gca ggt ctt gcc acc ttg gta gct tgc tcc tgg tat ggc cat	852
Ile Val Ala Gly Leu Ala Thr Leu Val Ala Cys Ser Trp Tyr Gly His	
130 135 140	
cag att gtc aca gac ttt tat aac cct ttg atc cct acc aac att aag	900
Gln Ile Val Thr Asp Phe Tyr Asn Pro Leu Ile Pro Thr Asn Ile Lys	
145 150 155	
tat gag ttt ggc cct gcc atc ttt att ggc tgg gca ggg tct gcc cta	948
Tyr Glu Phe Gly Pro Ala Ile Phe Ile Gly Trp Ala Gly Ser Ala Leu	
160 165 170	
gtc atc ctg gga ggt gca ctg ctc tcc tgt tcc tgt cct ggg aat gag	996
Val Ile Leu Gly Gly Ala Leu Leu Ser Cys Ser Cys Pro Gly Asn Glu	
175 180 185 190	
agc aag gct ggg tac cgt gca ccc cgc tct tac cct aag tcc aac tct	1044
Ser Lys Ala Gly Tyr Arg Ala Pro Arg Ser Tyr Pro Lys Ser Asn Ser	
195 200 205	
tcc aag gag tat gtg tgacctggga tctccttgcc ccagcctgac aggctatggg	1099
Ser Lys Glu Tyr Val	
210	
agtgtctaga tgccctgaaag ggcctggggc tgagctcagc ctgtgggcag ggtgccggac	1159
aaaggcctcc tggctactct gtccctgcac tccatgtata gtccctcttg gttgggggtg	1219
gggggggtgcc gttggtggga gagacaaaaa gagggagagt gtgctttttg tacagtaata	1279
aaaaataagt attgggaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa	1332

&lt;210&gt; 22

&lt;211&gt; 211

&lt;212&gt; PRT

&lt;213&gt; NM\_001307 claudin 7, CLDN7

&lt;400&gt; 22

Met Ala Asn Ser Gly Leu Gln Leu Leu Gly Phe Ser Met Ala Leu Leu  
1 5 10 15

Gly Trp Val Gly Leu Val Ala Cys Thr Ala Ile Pro Gln Trp Gln Met  
20 25 30

Ser Ser Tyr Ala Gly Asp Asn Ile Ile Thr Ala Gln Ala Met Tyr Lys  
35 40 45

Gly Leu Trp Met Asp Cys Val Thr Gln Ser Thr Gly Met Met Ser Cys  
50 55 60

Lys Met Tyr Asp Ser Val Leu Ala Leu Ser Ala Ala Leu Gln Ala Thr  
65 70 75 80

Arg Ala Leu Met Val Val Ser Leu Val Leu Gly Phe Leu Ala Met Phe  
85 90 95

Val Ala Thr Met Gly Met Lys Cys Thr Arg Cys Gly Gly Asp Asp Lys  
100 105 110

Val Lys Lys Ala Arg Ile Ala Met Gly Gly Gly Ile Ile Phe Ile Val  
115 120 125

Ala Gly Leu Ala Thr Leu Val Ala Cys Ser Trp Tyr Gly His Gln Ile  
130 135 140

Val Thr Asp Phe Tyr Asn Pro Leu Ile Pro Thr Asn Ile Lys Tyr Glu  
145 150 155 160

Phe Gly Pro Ala Ile Phe Ile Gly Trp Ala Gly Ser Ala Leu Val Ile  
165 170 175

Leu Gly Gly Ala Leu Leu Ser Cys Ser Cys Pro Gly Asn Glu Ser Lys  
180 185 190

Ala Gly Tyr Arg Ala Pro Arg Ser Tyr Pro Lys Ser Asn Ser Ser Lys  
195 200 205

Glu Tyr Val  
210

&lt;210&gt; 23

&lt;211&gt; 888

&lt;212&gt; DNA

&lt;213&gt; NM\_014736 KIAA0101 gene product

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (75)..(407)

&lt;223&gt;

&lt;400&gt; 23

ggggccttgt ccagtgaac accctcggct gggaagtcag ttctgttctct cctctcctct 60  
cttcttgttt gaac atg gtg cgg act aaa gca gac agt gtt cca ggc act 110  
Met Val Arg Thr Lys Ala Asp Ser Val Pro Gly Thr  
1 5 10  
tac aga aaa gtg gtg gct gct cga gcc ccc aga aag gtg ctt ggt tct 158  
Tyr Arg Lys Val Val Ala Ala Arg Ala Pro Arg Lys Val Leu Gly Ser  
15 20 25  
tcc acc tct gcc act aat tgc aca tca gtt tca tgc agg aaa gct gaa 206  
Ser Thr Ser Ala Thr Asn Ser Thr Ser Val Ser Ser Arg Lys Ala Glu  
30 35 40  
aat aaa tat gca gga ggg aac ccc gtt tgc gtg cgc cca act ccc aag 254  
Asn Lys Tyr Ala Gly Gly Asn Pro Val Cys Val Arg Pro Thr Pro Lys  
45 50 55 60  
tgg caa aaa gga att gga gaa ttc ttt agt ttg tcc cct aaa gat tct 302  
Trp Gln Lys Gly Ile Gly Glu Phe Phe Arg Leu Ser Pro Lys Asp Ser  
65 70 75  
gaa aaa gag aat cag att cct gaa gag gca gga agc agt ggc tta gga 350  
Glu Lys Glu Asn Gln Ile Pro Glu Glu Ala Gly Ser Ser Gly Leu Gly  
80 85 90  
aaa gca aag aga aaa gca tgt cct ttg caa cct gat cac aca aat gat 398  
Lys Ala Lys Arg Lys Ala Cys Pro Leu Gln Pro Asp His Thr Asn Asp  
95 100 105  
gaa aaa gaa tagaactttc tcattcatct ttgaataacg tctccttgtt 447  
Glu Lys Glu  
110  
taccctggta ttctagaatg taaatttaca taaatgtgtt tgttccaatt agctttgttg 507  
aacaggcatt taattaaataa atttaggttt aaatttagat gttcaaaagt agttgtgaaa 567  
tttgagaatt tgtaagacta attatggtaa cttagcttag tattcaatat aatgcattgt 627  
ttggtttctt ttaccaaatt aagtgtctag ttcttgctaa aatcaagtca ttgcattgtg 687  
ttctaattac aagtatgttg tatttgagat ttgcttagat tgttgtactg ctgccatttt 747  
tattggtgtt tgattattgg aatggtgcc aattgtcact ccttctactt gctttaaaaa 807  
gcagagttag atttttgcac attaaaaaaa ttcagtatta attaaacact aaaaaaaaaa 867

aaaaaaaaa aaaaaaaaaa a

888

&lt;210&gt; 24

&lt;211&gt; 111

&lt;212&gt; PRT

&lt;213&gt; NM\_014736 KIAA0101 gene product

&lt;400&gt; 24

Met Val Arg Thr Lys Ala Asp Ser Val Pro Gly Thr Tyr Arg Lys Val  
 1 5 10 15

Val Ala Ala Arg Ala Pro Arg Lys Val Leu Gly Ser Ser Thr Ser Ala  
 20 25 30

Thr Asn Ser Thr Ser Val Ser Ser Arg Lys Ala Glu Asn Lys Tyr Ala  
 35 40 45

Gly Gly Asn Pro Val Cys Val Arg Pro Thr Pro Lys Trp Gln Lys Gly  
 50 55 60

Ile Gly Glu Phe Phe Arg Leu Ser Pro Lys Asp Ser Glu Lys Glu Asn  
 65 70 75 80

Gln Ile Pro Glu Glu Ala Gly Ser Ser Gly Leu Gly Lys Ala Lys Arg  
 85 90 95

Lys Ala Cys Pro Leu Gln Pro Asp His Thr Asn Asp Glu Lys Glu  
 100 105 110

&lt;210&gt; 25

&lt;211&gt; 598

&lt;212&gt; DNA

&lt;213&gt; NM\_003064 secretory leukocyte protease inhibitor, SLPI

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (23)..(418)

&lt;223&gt;

&lt;400&gt; 25

cagagtcact cctgccttca cc atg aag tcc agc ggc ctc ttc ccc ttc ctg  
 Met Lys Ser Ser Gly Leu Phe Pro Phe Leu  
 1 5 10

52



gtg ctg ctt gcc ctg gga act ctg gca cct tgg gct gtg gaa ggc tct 100  
 Val Leu Leu Ala Leu Gly Thr Leu Ala Pro Trp Ala Val Glu Gly Ser  
 15 20 25

gga aag tcc ttc aaa gct gga gtc tgt cct cct aag aaa tct gcc cag 148  
 Gly Lys Ser Phe Lys Ala Gly Val Cys Pro Pro Lys Lys Ser Ala Gln  
 30 35 40

tgc ctt aga tac aag aaa cct gag tgc cag agt gac tgg cag tgt cca 196  
 Cys Leu Arg Tyr Lys Lys Pro Glu Cys Gln Ser Asp Trp Gln Cys Pro  
 45 50 55

ggg aag aag aga tgt tgt cct gac act tgt ggc atc aaa tgc ctg gat 244  
 Gly Lys Lys Arg Cys Cys Pro Asp Thr Cys Gly Ile Lys Cys Leu Asp  
 60 65 70

cct gtt gac acc cca aac cca aca agg agg aag cct ggg aag tgc cca 292  
 Pro Val Asp Thr Pro Asn Pro Thr Arg Arg Lys Pro Gly Lys Cys Pro  
 75 80 85 90

gtg act tat ggc caa tgt ttg atg ctt aac ccc ccc aat ttc tgt gag 340  
 Val Thr Tyr Gly Gln Cys Leu Met Leu Asn Pro Pro Asn Phe Cys Glu  
 95 100 105

atg gat ggc cag tgc aag cgt gac ttg aag tgt tgc atg ggc atg tgt 388  
 Met Asp Gly Gln Cys Lys Arg Asp Leu Lys Cys Cys Met Gly Met Cys  
 110 115 120

ggg aaa tcc tgc gtt tcc cct gtg aaa gct tgattcctgc catatggagg 438  
 Gly Lys Ser Cys Val Ser Pro Val Lys Ala  
 125 130

aggctctgga gtctgtctct gtgtggtcca ggtcctttcc accctgagac ttggctccac 498

cactgatatc ctcttttggg gaaaggcttg gcacacagca ggctttcaag aagtgccagt 558

tgatcaatga ataaataaac gagcctattt ctctttgcac 598

<210> 26

<211> 132

<212> PRT

<213> NM\_003064 secretory leukocyte protease inhibitor, SLPI

<400> 26

Met Lys Ser Ser Gly Leu Phe Pro Phe Leu Val Leu Leu Ala Leu Gly  
 1 5 10 15

Thr Leu Ala Pro Trp Ala Val Glu Gly Ser Gly Lys Ser Phe Lys Ala  
 20 25 30

Gly Val Cys Pro Pro Lys Lys Ser Ala Gln Cys Leu Arg Tyr Lys Lys  
 35 40 45

Pro Glu Cys Gln Ser Asp Trp Gln Cys Pro Gly Lys Lys Arg Cys Cys  
 50 55 60

Pro Asp Thr Cys Gly Ile Lys Cys Leu Asp Pro Val Asp Thr Pro Asn  
65 70 75 80

Pro Thr Arg Arg Lys Pro Gly Lys Cys Pro Val Thr Tyr Gly Gln Cys  
85 90 95

Leu Met Leu Asn Pro Pro Asn Phe Cys Glu Met Asp Gly Gln Cys Lys  
100 105 110

Arg Asp Leu Lys Cys Cys Met Gly Met Cys Gly Lys Ser Cys Val Ser  
115 120 125

Pro Val Lys Ala  
130

<210> 27

<211> 3970

<212> DNA

<213> NM\_013994 DDR1

<220>

<221> CDS

<222> (337)..(3093)

<223>

<400> 27

ggcttaggaa gtattaactg atctctgccc tagttctcat gtgttaaata tggatagtaa 60

tagtatctac cttatgaagt gactgtgaag ataaaattat ggattctgtt taagggttta 120

ggccagtgtc tggcacaggg gaagcattct aaaaatatag ctgatgctgt taaacaatga 180

ctgttggtgt tgttttactg ttattatccc caaagcggcc cattctgtct gttgctgtca 240

gctatgactc agtcccctga ttaacttaag caccacccat tttatcccct gcagagatgc 300

tgccccacc cccttaggcc cgagggatca ggagct atg gga cca gag gcc ctg 354  
Met Gly Pro Glu Ala Leu  
1 5

tca tct tta ctg ctg ctg ctc ttg gtg gca agt gga gat gct gac atg 402  
Ser Ser Leu Leu Leu Leu Leu Val Ala Ser Gly Asp Ala Asp Met  
10 15 20

aag gga cat ttt gat cct gcc aag tgc cgc tat gcc ctg ggc atg cag 450  
Lys Gly His Phe Asp Pro Ala Lys Cys Arg Tyr Ala Leu Gly Met Gln  
25 30 35

gac cgg acc atc cca gac agt gac atc tct gct tcc agc tcc tgg tca 498  
Asp Arg Thr Ile Pro Asp Ser Asp Ile Ser Ala Ser Ser Ser Trp Ser  
40 45 50

gat tcc act gcc gcc cgc cac agc agg ttg gag agc agt gac ggg gat 546

101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840. 841. 842. 843. 844. 845. 846. 847. 848. 849. 850. 851. 852. 853. 854. 855. 856. 857. 858. 859. 860. 861. 862. 863. 864. 865. 866. 867. 868. 869. 870. 871. 872. 873. 874. 875. 876. 877. 878. 879. 880. 881. 882. 883. 884. 885. 886. 887. 888. 889. 890. 891. 892. 893. 894. 895. 896. 897. 898. 899. 900. 901. 902. 903. 904. 905. 906. 907. 908. 909. 910. 911. 912. 913. 914. 915. 916. 917. 918. 919.

|   |      |
|---|------|
| ggg gac ccc aga gcc cgg gct gtc tca gtg ccc ctt ggc ggc cgt gtg<br>Gly Asp Pro Arg Ala Arg Ala Val Ser Val Pro Leu Gly Gly Arg Val<br>330 335 340     | 1362 |
| gct cgc ttt ctg cag tgc cgc ttc ctc ttt gcg ggg ccc tgg tta ctc<br>Ala Arg Phe Leu Gln Cys Arg Phe Leu Phe Ala Gly Pro Trp Leu Leu<br>345 350 355     | 1410 |
| ttc agc gaa atc tcc ttc atc tct gat gtg gtg aac aat tcc tct ccg<br>Phe Ser Glu Ile Ser Phe Ile Ser Asp Val Val Asn Asn Ser Ser Pro<br>360 365 370     | 1458 |
| gca ctg gga ggc acc ttc ccg cca gcc ccc tgg tgg ccg cct ggc cca<br>Ala Leu Gly Gly Thr Phe Pro Pro Ala Pro Trp Trp Pro Pro Gly Pro<br>375 380 385 390 | 1506 |
| cct ccc acc aac ttc agc agc ttg gag ctg gag ccc aga ggc cag cag<br>Pro Pro Thr Asn Phe Ser Ser Leu Glu Leu Glu Pro Arg Gly Gln Gln<br>395 400 405     | 1554 |
| ccc gtg gcc aag gcc gag ggg agc ccg acc gcc atc ctc atc ggc tgc<br>Pro Val Ala Lys Ala Glu Gly Ser Pro Thr Ala Ile Leu Ile Gly Cys<br>410 415 420     | 1602 |
| ctg gtg gcc atc atc ctg ctc ctg ctg ctc atc att gcc ctc atg ctc<br>Leu Val Ala Ile Ile Leu Leu Leu Leu Leu Ile Ile Ala Leu Met Leu<br>425 430 435     | 1650 |
| tgg cgg ctg cac tgg cgc agg ctc ctc agc aag gct gaa cgg agg gtg<br>Trp Arg Leu His Trp Arg Arg Leu Leu Ser Lys Ala Glu Arg Arg Val<br>440 445 450     | 1698 |
| ttg gaa gag gag ctg acg gtt cac ctc tct gtc cct ggg gac act atc<br>Leu Glu Glu Glu Leu Thr Val His Leu Ser Val Pro Gly Asp Thr Ile<br>455 460 465 470 | 1746 |
| ctc atc aac aac cgc cca ggt cct aga gag cca ccc ccg tac cag gag<br>Leu Ile Asn Asn Arg Pro Gly Pro Arg Glu Pro Pro Pro Tyr Gln Glu<br>475 480 485     | 1794 |
| ccc cgg cct cgt ggg aat ccg ccc cac tcc gct ccc tgt gtc ccc aat<br>Pro Arg Pro Arg Gly Asn Pro Pro His Ser Ala Pro Cys Val Pro Asn<br>490 495 500     | 1842 |
| ggc tct gcg ttg ctg ctc tcc aat cca gcc tac cgc ctc ctt ctg gcc<br>Gly Ser Ala Leu Leu Leu Ser Asn Pro Ala Tyr Arg Leu Leu Leu Ala<br>505 510 515     | 1890 |
| act tac gcc cgt ccc cct cga ggc ccg ggc ccc ccc aca ccc gcc tgg<br>Thr Tyr Ala Arg Pro Pro Arg Gly Pro Gly Pro Pro Thr Pro Ala Trp<br>520 525 530     | 1938 |
| gcc aaa ccc acc aac acc cag gcc tac agt ggg gac tat atg gag cct<br>Ala Lys Pro Thr Asn Thr Gln Ala Tyr Ser Gly Asp Tyr Met Glu Pro<br>535 540 545 550 | 1986 |
| gag aag cca ggc gcc ccg ctt ctg ccc cca cct ccc cag aac agc gtc<br>Glu Lys Pro Gly Ala Pro Leu Leu Pro Pro Pro Pro Gln Asn Ser Val<br>555 560 565     | 2034 |
| ccc cat tat gcc gag gct gac att gtt acc ctg cag ggc gtc acc ggg<br>Pro His Tyr Ala Glu Ala Asp Ile Val Thr Leu Gln Gly Val Thr Gly<br>570 575 580     | 2082 |

|   |      |
|---|------|
| ggc aac acc tat gct gtg cct gca ctg ccc cca ggg gca gtc ggg gat<br>Gly Asn Thr Tyr Ala Val Pro Ala Leu Pro Pro Gly Ala Val Gly Asp<br>585 590 595     | 2130 |
| ggg ccc ccc aga gtg gat ttc cct cga tct cga ctc cgc ttc aag gag<br>Gly Pro Pro Arg Val Asp Phe Pro Arg Ser Arg Leu Arg Phe Lys Glu<br>600 605 610     | 2178 |
| aag ctt ggc gag ggc cag ttt ggg gag gtg cac ctg tgt gag gtc gac<br>Lys Leu Gly Glu Gly Gln Phe Gly Glu Val His Leu Cys Glu Val Asp<br>615 620 625 630 | 2226 |
| agc cct caa gat ctg gtt agt ctt gat ttc ccc ctt aat gtg cgt aag<br>Ser Pro Gln Asp Leu Val Ser Leu Asp Phe Pro Leu Asn Val Arg Lys<br>635 640 645     | 2274 |
| gga cac cct ttg ctg gta gct gtc aag atc tta cgg cca gat gcc acc<br>Gly His Pro Leu Leu Val Ala Val Lys Ile Leu Arg Pro Asp Ala Thr<br>650 655 660     | 2322 |
| aag aat gcc agc ttc tcc ttg ttc tcc agg aat gat ttc ctg aaa gag<br>Lys Asn Ala Ser Phe Ser Leu Phe Ser Arg Asn Asp Phe Leu Lys Glu<br>665 670 675     | 2370 |
| gtg aag atc atg tcg agg ctc aag gac cca aac atc att cgg ctg ctg<br>Val Lys Ile Met Ser Arg Leu Lys Asp Pro Asn Ile Ile Arg Leu Leu<br>680 685 690     | 2418 |
| ggc gtg tgt gtg cag gac gac ccc ctc tgc atg att act gac tac atg<br>Gly Val Cys Val Gln Asp Pro Leu Cys Met Ile Thr Asp Tyr Met<br>695 700 705 710     | 2466 |
| gag aac ggc gac ctc aac cag ttc ctc agt gcc cac cag ctg gag gac<br>Glu Asn Gly Asp Leu Asn Gln Phe Leu Ser Ala His Gln Leu Glu Asp<br>715 720 725     | 2514 |
| aag gca gcc gag ggg gcc cct ggg gac ggg cag gct gcg cag ggg ccc<br>Lys Ala Ala Glu Gly Ala Pro Gly Asp Gly Gln Ala Ala Gln Gly Pro<br>730 735 740     | 2562 |
| acc atc agc tac cca atg ctg ctg cat gtg gca gcc cag atc gcc tcc<br>Thr Ile Ser Tyr Pro Met Leu Leu His Val Ala Ala Gln Ile Ala Ser<br>745 750 755     | 2610 |
| ggc atg cgc tat ctg gcc aca ctc aac ttt gta cat cgg gac ctg gcc<br>Gly Met Arg Tyr Leu Ala Thr Leu Asn Phe Val His Arg Asp Leu Ala<br>760 765 770     | 2658 |
| acg cgg aac tgc cta gtt ggg gaa aat ttc acc atc aaa atc gca gac<br>Thr Arg Asn Cys Leu Val Gly Glu Asn Phe Thr Ile Lys Ile Ala Asp<br>775 780 785 790 | 2706 |
| ttt ggc atg agc cgg aac ctc tat gct ggg gac tat tac cgt gtg cag<br>Phe Gly Met Ser Arg Asn Leu Tyr Ala Gly Asp Tyr Tyr Arg Val Gln<br>795 800 805     | 2754 |
| ggc cgg gca gtg ctg ccc atc cgc tgg atg gcc tgg gag tgc atc ctc<br>Gly Arg Ala Val Leu Pro Ile Arg Trp Met Ala Trp Glu Cys Ile Leu<br>810 815 820     | 2802 |
| atg ggg aag ttc acg act gcg agt gac gtg tgg gcc ttt ggt gtg acc<br>Met Gly Lys Phe Thr Thr Ala Ser Asp Val Trp Ala Phe Gly Val Thr<br>825 830 835     | 2850 |
| ctg tgg gag gtg ctg atg ctc tgt agg gcc cag ccc ttt ggg tca gct<br>Leu Trp Glu Val Leu Met Leu Cys Arg Ala Gln Pro Phe Gly Ser Ala                    | 2898 |

| 840   | 845 | 850 |      |
|---|-----|-----|------|
| cac cga cga gca ggt cat cga gaa cgc ggg gga gtt ctt ccg gga cca   |     |     | 2946 |
| His Arg Arg Ala Gly His Arg Glu Arg Gly Gly Val Leu Pro Gly Pro   |     |     |      |
| 855   | 860 | 865 | 870  |
| ggg ccg gca gtg tac ctg tcc cgg ccg cct gcc tgc ccg cag ggc cta   |     |     | 2994 |
| Gly Pro Ala Val Tyr Leu Ser Arg Pro Pro Ala Cys Pro Gln Gly Leu   |     |     |      |
|   | 875 | 880 | 885  |
| tat gag ctg atg ctt cgg tgc tgg agc cgg gag tct gag cag cga cca   |     |     | 3042 |
| Tyr Glu Leu Met Leu Arg Cys Trp Ser Arg Glu Ser Glu Gln Arg Pro   |     |     |      |
|   | 890 | 895 | 900  |
| ccc ttt tcc cag ctg cat cgg ttc ctg gca gag gat gca ctc aac acg   |     |     | 3090 |
| Pro Phe Ser Gln Leu His Arg Phe Leu Ala Glu Asp Ala Leu Asn Thr   |     |     |      |
|   | 905 | 910 | 915  |
| gtg tgaatcacac atccagctgc ccctccctca gggagcgcac cagggaaga         |     |     | 3143 |
| Val   |     |     |      |
| cagtgcact aaaacaagag gacacaatgg cacctctgcc ctccccctcc cgacagccca  |     |     | 3203 |
| tcacctctaa tagaggcagt gagactgcag gtgggctggg cccaccagg gagctgatgc  |     |     | 3263 |
| cccttctccc ctctctggac acactctcat gtccccctcc tgttcttctt tctagaagc  |     |     | 3323 |
| ccccctgtcg cccaccaggc tggctctgtg gatgggatcc tctccaccct cctctagcca |     |     | 3383 |
| tcccttgagg aagggtgggg agaaatatag gatagacact ggacatggcc cattggagca |     |     | 3443 |
| cctgggcccc actggacaac actgattcct ggagagggtg ctgcgcccc agcttctctc  |     |     | 3503 |
| tccctgtcac aactggacc cactggctg agaactctgg ggtgaggagg acaagaagga   |     |     | 3563 |
| gaggaaaatg tttccttggt cctgctcctg tacttgtctt cagcttgggc ttcttctctc |     |     | 3623 |
| tccatcacct gaaacactgg acctgggggt agccccgcc cagccctcag tcacccccac  |     |     | 3683 |
| ttcccacttg cagtcttgta gctagaactt ctctaagcct atacgtttct gtggagtaaa |     |     | 3743 |
| tattgggatt ggggggaaag agggagcaac ggcccatagc cttgggggtg gacatctcta |     |     | 3803 |
| gtgtagctgc cacattgatt tttctataat cacttggggt ttgtacattt ttggggggag |     |     | 3863 |
| agacacagat ttttaccta atatatggac ctagcttgag gcaattttaa tcccctgcac  |     |     | 3923 |
| taggcaggta ataataaagg ttgagttttc cacaaaaaa aaaaaaa                |     |     | 3970 |

&lt;210&gt; 28

&lt;211&gt; 919

&lt;212&gt; PRT

&lt;213&gt; NM\_013994 DDR1

&lt;400&gt; 28

Met Gly Pro Glu Ala Leu Ser Ser Leu Leu Leu Leu Leu Val Ala  
 1 5 10 15

Ser Gly Asp Ala Asp Met Lys Gly His Phe Asp Pro Ala Lys Cys Arg

20 25 30

Tyr Ala Leu Gly Met Gln Asp Arg Thr Ile Pro Asp Ser Asp Ile Ser  
35 40 45

Ala Ser Ser Ser Trp Ser Asp Ser Thr Ala Ala Arg His Ser Arg Leu  
50 55 60

Glu Ser Ser Asp Gly Asp Gly Ala Trp Cys Pro Ala Gly Ser Val Phe  
65 70 75 80

Pro Lys Glu Glu Glu Tyr Leu Gln Val Asp Leu Gln Arg Leu His Leu  
85 90 95

Val Ala Leu Val Gly Thr Gln Gly Arg His Ala Gly Gly Leu Gly Lys  
100 105 110

Glu Phe Ser Arg Ser Tyr Arg Leu Arg Tyr Ser Arg Asp Gly Arg Arg  
115 120 125

Trp Met Gly Trp Lys Asp Arg Trp Gly Gln Glu Val Ile Ser Gly Asn  
130 135 140

Glu Asp Pro Glu Gly Val Val Leu Lys Asp Leu Gly Pro Pro Met Val  
145 150 155 160

Ala Arg Leu Val Arg Phe Tyr Pro Arg Ala Asp Arg Val Met Ser Val  
165 170 175

Cys Leu Arg Val Glu Leu Tyr Gly Cys Leu Trp Arg Asp Gly Leu Leu  
180 185 190

Ser Tyr Thr Ala Pro Val Gly Gln Thr Met Tyr Leu Ser Glu Ala Val  
195 200 205

Tyr Leu Asn Asp Ser Thr Tyr Asp Gly His Thr Val Gly Gly Leu Gln  
210 215 220

Tyr Gly Gly Leu Gly Gln Leu Ala Asp Gly Val Val Gly Leu Asp Asp  
225 230 235 240

Phe Arg Lys Ser Gln Glu Leu Arg Val Trp Pro Gly Tyr Asp Tyr Val  
245 250 255

Gly Trp Ser Asn His Ser Phe Ser Ser Gly Tyr Val Glu Met Glu Phe  
260 265 270

Glu Phe Asp Arg Leu Arg Ala Phe Gln Ala Met Gln Val His Cys Asn  
275 280 285

Asn Met His Thr Leu Gly Ala Arg Leu Pro Gly Gly Val Glu Cys Arg

290                      295                      300  
 Phe Arg Arg Gly Pro Ala Met Ala Trp Glu Gly Glu Pro Met Arg His  
 305                      310                      315                      320  
 Asn Leu Gly Gly Asn Leu Gly Asp Pro Arg Ala Arg Ala Val Ser Val  
 325                      330                      335  
 Pro Leu Gly Gly Arg Val Ala Arg Phe Leu Gln Cys Arg Phe Leu Phe  
 340                      345                      350  
 Ala Gly Pro Trp Leu Leu Phe Ser Glu Ile Ser Phe Ile Ser Asp Val  
 355                      360                      365  
 Val Asn Asn Ser Ser Pro Ala Leu Gly Gly Thr Phe Pro Pro Ala Pro  
 370                      375                      380  
 Trp Trp Pro Pro Gly Pro Pro Pro Thr Asn Phe Ser Ser Leu Glu Leu  
 385                      390                      395                      400  
 Glu Pro Arg Gly Gln Gln Pro Val Ala Lys Ala Glu Gly Ser Pro Thr  
 405                      410                      415  
 Ala Ile Leu Ile Gly Cys Leu Val Ala Ile Ile Leu Leu Leu Leu Leu  
 420                      425                      430  
 Ile Ile Ala Leu Met Leu Trp Arg Leu His Trp Arg Arg Leu Leu Ser  
 435                      440                      445  
 Lys Ala Glu Arg Arg Val Leu Glu Glu Glu Leu Thr Val His Leu Ser  
 450                      455                      460  
 Val Pro Gly Asp Thr Ile Leu Ile Asn Asn Arg Pro Gly Pro Arg Glu  
 465                      470                      475                      480  
 Pro Pro Pro Tyr Gln Glu Pro Arg Pro Arg Gly Asn Pro Pro His Ser  
 485                      490                      495  
 Ala Pro Cys Val Pro Asn Gly Ser Ala Leu Leu Leu Ser Asn Pro Ala  
 500                      505                      510  
 Tyr Arg Leu Leu Leu Ala Thr Tyr Ala Arg Pro Pro Arg Gly Pro Gly  
 515                      520                      525  
 Pro Pro Thr Pro Ala Trp Ala Lys Pro Thr Asn Thr Gln Ala Tyr Ser  
 530                      535                      540  
 Gly Asp Tyr Met Glu Pro Glu Lys Pro Gly Ala Pro Leu Leu Pro Pro  
 545                      550                      555                      560



Pro Pro Gln Asn Ser Val Pro His Tyr Ala Glu Ala Asp Ile Val Thr  
565 570 575

Leu Gln Gly Val Thr Gly Gly Asn Thr Tyr Ala Val Pro Ala Leu Pro  
580 585 590

Pro Gly Ala Val Gly Asp Gly Pro Pro Arg Val Asp Phe Pro Arg Ser  
595 600 605

Arg Leu Arg Phe Lys Glu Lys Leu Gly Glu Gly Gln Phe Gly Glu Val  
610 615 620

His Leu Cys Glu Val Asp Ser Pro Gln Asp Leu Val Ser Leu Asp Phe  
625 630 635 640

Pro Leu Asn Val Arg Lys Gly His Pro Leu Leu Val Ala Val Lys Ile  
645 650 655

Leu Arg Pro Asp Ala Thr Lys Asn Ala Ser Phe Ser Leu Phe Ser Arg  
660 665 670

Asn Asp Phe Leu Lys Glu Val Lys Ile Met Ser Arg Leu Lys Asp Pro  
675 680 685

Asn Ile Ile Arg Leu Leu Gly Val Cys Val Gln Asp Asp Pro Leu Cys  
690 695 700

Met Ile Thr Asp Tyr Met Glu Asn Gly Asp Leu Asn Gln Phe Leu Ser  
705 710 715 720

Ala His Gln Leu Glu Asp Lys Ala Ala Glu Gly Ala Pro Gly Asp Gly  
725 730 735

Gln Ala Ala Gln Gly Pro Thr Ile Ser Tyr Pro Met Leu Leu His Val  
740 745 750

Ala Ala Gln Ile Ala Ser Gly Met Arg Tyr Leu Ala Thr Leu Asn Phe  
755 760 765

Val His Arg Asp Leu Ala Thr Arg Asn Cys Leu Val Gly Glu Asn Phe  
770 775 780

Thr Ile Lys Ile Ala Asp Phe Gly Met Ser Arg Asn Leu Tyr Ala Gly  
785 790 795 800

Asp Tyr Tyr Arg Val Gln Gly Arg Ala Val Leu Pro Ile Arg Trp Met  
805 810 815

Ala Trp Glu Cys Ile Leu Met Gly Lys Phe Thr Thr Ala Ser Asp Val  
820 825 830

Trp Ala Phe Gly Val Thr Leu Trp Glu Val Leu Met Leu Cys Arg Ala  
835 840 845

Gln Pro Phe Gly Ser Ala His Arg Arg Ala Gly His Arg Glu Arg Gly  
850 855 860

Gly Val Leu Pro Gly Pro Gly Pro Ala Val Tyr Leu Ser Arg Pro Pro  
865 870 875 880

Ala Cys Pro Gln Gly Leu Tyr Glu Leu Met Leu Arg Cys Trp Ser Arg  
885 890 895

Glu Ser Glu Gln Arg Pro Pro Phe Ser Gln Leu His Arg Phe Leu Ala  
900 905 910

Glu Asp Ala Leu Asn Thr Val  
915

<210> 29

<211> 5698

<212> DNA

<213> NM\_001067 TOP2A

<220>

<221> CDS

<222> (127)..(4719)

<223>

<400> 29

agggttcaagt ggagctctcc taaccgacgc gcgtctgtgg agaagcggct tggtcggggg 60

tggctctcgtg gggtcctgcc tggttagtcg ctttcagggt tcttgagccc cttcagcacc 120

gtcacc atg gaa gtg tca cca ttg cag cct gta aat gaa aat atg caa 168  
Met Glu Val Ser Pro Leu Gln Pro Val Asn Glu Asn Met Gln  
1 5 10

gtc aac aaa ata aag aaa aat gaa gat gct aag aaa aga ctg tct gtt 216  
Val Asn Lys Ile Lys Lys Asn Glu Asp Ala Lys Lys Arg Leu Ser Val  
15 20 25 30

gaa aga atc tat caa aag aaa aca caa ttg gaa cat att ttg ctc cgc 264  
Glu Arg Ile Tyr Gln Lys Lys Thr Gln Leu Glu His Ile Leu Leu Arg  
35 40 45

cca gac acc tac att ggt tct gtg gaa tta gtg acc cag caa atg tgg 312  
Pro Asp Thr Tyr Ile Gly Ser Val Glu Leu Val Thr Gln Gln Met Trp  
50 55 60

gtt tac gat gaa gat gtt ggc att aac tat agg gaa gtc act ttt gtt 360

|   |      |
|---|------|
| Val Tyr Asp Glu Asp Val Gly Ile Asn Tyr Arg Glu Val Thr Phe Val |      |
| 65 70 75  |      |
| cct ggt ttg tac aaa atc ttt gat gag att cta gtt aat gct gcg gac | 408  |
| Pro Gly Leu Tyr Lys Ile Phe Asp Glu Ile Leu Val Asn Ala Ala Asp |      |
| 80 85 90  |      |
| aac aaa caa agg gac cca aaa atg tct tgt att aga gtc aca att gat | 456  |
| Asn Lys Gln Arg Asp Pro Lys Met Ser Cys Ile Arg Val Thr Ile Asp |      |
| 95 100 105 110  |      |
| ccg gaa aac aat tta att agt ata tgg aat aat gga aaa ggt att cct | 504  |
| Pro Glu Asn Asn Leu Ile Ser Ile Trp Asn Asn Gly Lys Gly Ile Pro |      |
| 115 120 125   |      |
| gtt gtt gaa cac aaa gtt gaa aag atg tat gtc cca gct ctc ata ttt | 552  |
| Val Val Glu His Lys Val Glu Lys Met Tyr Val Pro Ala Leu Ile Phe |      |
| 130 135 140   |      |
| gga cag ctc cta act tct agt aac tat gat gat gat gaa aag aaa gtg | 600  |
| Gly Gln Leu Leu Thr Ser Ser Asn Tyr Asp Asp Asp Glu Lys Lys Val |      |
| 145 150 155   |      |
| aca ggt ggt cga aat ggc tat gga gcc aaa ttg tgt aac ata ttc agt | 648  |
| Thr Gly Gly Arg Asn Gly Tyr Gly Ala Lys Leu Cys Asn Ile Phe Ser |      |
| 160 165 170   |      |
| acc aaa ttt act gtg gaa aca gcc agt aga gaa tac aag aaa atg ttc | 696  |
| Thr Lys Phe Thr Val Glu Thr Ala Ser Arg Glu Tyr Lys Lys Met Phe |      |
| 175 180 185 190   |      |
| aaa cag aca tgg atg gat aat atg gga aga gct ggt gag atg gaa ctc | 744  |
| Lys Gln Thr Trp Met Asp Asn Met Gly Arg Ala Gly Glu Met Glu Leu |      |
| 195 200 205   |      |
| aag ccc ttc aat gga gaa gat tat aca tgt atc acc ttt cag cct gat | 792  |
| Lys Pro Phe Asn Gly Glu Asp Tyr Thr Cys Ile Thr Phe Gln Pro Asp |      |
| 210 215 220   |      |
| ttg tct aag ttt aaa atg caa agc ctg gac aaa gat att gtt gca cta | 840  |
| Leu Ser Lys Phe Lys Met Gln Ser Leu Asp Lys Asp Ile Val Ala Leu |      |
| 225 230 235   |      |
| atg gtc aga aga gca tat gat att gct gga tcc acc aaa gat gtc aaa | 888  |
| Met Val Arg Arg Ala Tyr Asp Ile Ala Gly Ser Thr Lys Asp Val Lys |      |
| 240 245 250   |      |
| gtc ttt ctt aat gga aat aaa ctg cca gta aaa gga ttt cgt agt tat | 936  |
| Val Phe Leu Asn Gly Asn Lys Leu Pro Val Lys Gly Phe Arg Ser Tyr |      |
| 255 260 265 270   |      |
| gtg gac atg tat ttg aag gac aag ttg gat gaa act ggt aac tcc ttg | 984  |
| Val Asp Met Tyr Leu Lys Asp Lys Leu Asp Glu Thr Gly Asn Ser Leu |      |
| 275 280 285   |      |
| aaa gta ata cat gaa caa gta aac cac agg tgg gaa gtg tgt tta act | 1032 |
| Lys Val Ile His Glu Gln Val Asn His Arg Trp Glu Val Cys Leu Thr |      |
| 290 295 300   |      |
| atg agt gaa aaa ggc ttt cag caa att agc ttt gtc aac agc att gct | 1080 |
| Met Ser Glu Lys Gly Phe Gln Gln Ile Ser Phe Val Asn Ser Ile Ala |      |
| 305 310 315   |      |
| aca tcc aag ggt ggc aga cat gtt gat tat gta gct gat cag att gtg | 1128 |
| Thr Ser Lys Gly Gly Arg His Val Asp Tyr Val Ala Asp Gln Ile Val |      |
| 320 325 330   |      |

|   |      |
|---|------|
| act aaa ctt gtt gat gtt gtg aag aag aag aac aag ggt ggt gtt gca<br>Thr Lys Leu Val Asp Val Val Lys Lys Lys Asn Lys Gly Gly Val Ala<br>335 340 345 350 | 1176 |
| gta aaa gca cat cag gtg aaa aat cac atg tgg att ttt gta aat gcc<br>Val Lys Ala His Gln Val Lys Asn His Met Trp Ile Phe Val Asn Ala<br>355 360 365     | 1224 |
| tta att gaa aac cca acc ttt gac tct cag aca aaa gaa aac atg act<br>Leu Ile Glu Asn Pro Thr Phe Asp Ser Gln Thr Lys Glu Asn Met Thr<br>370 375 380     | 1272 |
| tta caa ccc aag agc ttt gga tca aca tgc caa ttg agt gaa aaa ttt<br>Leu Gln Pro Lys Ser Phe Gly Ser Thr Cys Gln Leu Ser Glu Lys Phe<br>385 390 395     | 1320 |
| atc aaa gct gcc att ggc tgt ggt att gta gaa agc ata cta aac tgg<br>Ile Lys Ala Ala Ile Gly Cys Gly Ile Val Glu Ser Ile Leu Asn Trp<br>400 405 410     | 1368 |
| gtg aag ttt aag gcc caa gtc cag tta aac aag aag tgt tca gct gta<br>Val Lys Phe Lys Ala Gln Val Gln Leu Asn Lys Lys Cys Ser Ala Val<br>415 420 425 430 | 1416 |
| aaa cat aat aga atc aag gga att ccc aaa ctc gat gat gcc aat gat<br>Lys His Asn Arg Ile Lys Gly Ile Pro Lys Leu Asp Asp Ala Asn Asp<br>435 440 445     | 1464 |
| gca ggg ggc cga aac tcc act gag tgt acg ctt atc ctg act gag gga<br>Ala Gly Gly Arg Asn Ser Thr Glu Cys Thr Leu Ile Leu Thr Glu Gly<br>450 455 460     | 1512 |
| gat tca gcc aaa act ttg gct gtt tca ggc ctt ggt gtg gtt ggg aga<br>Asp Ser Ala Lys Thr Leu Ala Val Ser Gly Leu Gly Val Val Gly Arg<br>465 470 475     | 1560 |
| gac aaa tat ggg gtt ttc cct ctt aga gga aaa ata ctc aat gtt cga<br>Asp Lys Tyr Gly Val Phe Pro Leu Arg Gly Lys Ile Leu Asn Val Arg<br>480 485 490     | 1608 |
| gaa gct tct cat aag cag atc atg gaa aat gct gag att aac aat atc<br>Glu Ala Ser His Lys Gln Ile Met Glu Asn Ala Glu Ile Asn Asn Ile<br>495 500 505 510 | 1656 |
| atc aag att gtg ggt ctt cag tac aag aaa aac tat gaa gat gaa gat<br>Ile Lys Ile Val Gly Leu Gln Tyr Lys Lys Asn Tyr Glu Asp Glu Asp<br>515 520 525     | 1704 |
| tca ttg aag acg ctt cgt tat ggg aag ata atg att atg aca gat cag<br>Ser Leu Lys Thr Leu Arg Tyr Gly Lys Ile Met Ile Met Thr Asp Gln<br>530 535 540     | 1752 |
| gac caa gat ggt tcc cac atc aaa ggc ttg ctg att aat ttt atc cat<br>Asp Gln Asp Gly Ser His Ile Lys Gly Leu Leu Ile Asn Phe Ile His<br>545 550 555     | 1800 |
| cac aac tgg ccc tct ctt ctg cga cat cgt ttt ctg gag gaa ttt atc<br>His Asn Trp Pro Ser Leu Leu Arg His Arg Phe Leu Glu Glu Phe Ile<br>560 565 570     | 1848 |
| act ccc att gta aag gta tct aaa aac aag caa gaa atg gca ttt tac<br>Thr Pro Ile Val Lys Val Ser Lys Asn Lys Gln Glu Met Ala Phe Tyr<br>575 580 585 590 | 1896 |
| agc ctt cct gaa ttt gaa gag tgg aag agt tct act cca aat cat aaa   | 1944 |

|   |      |
|---|------|
| Ser Leu Pro Glu Phe Glu Glu Trp Lys Ser Ser Thr Pro Asn His Lys<br>595 600 605  |      |
| aaa tgg aaa gtc aaa tat tac aaa ggt ttg ggc acc agc aca tca aag<br>Lys Trp Lys Val Lys Tyr Tyr Lys Gly Leu Gly Thr Ser Thr Ser Lys<br>610 615 620     | 1992 |
| gaa gct aaa gaa tac ttt gca gat atg aaa aga cat cgt atc cag ttc<br>Glu Ala Lys Glu Tyr Phe Ala Asp Met Lys Arg His Arg Ile Gln Phe<br>625 630 635     | 2040 |
| aaa tat tct ggt cct gaa gat gat gct gct atc agc ctg gcc ttt agc<br>Lys Tyr Ser Gly Pro Glu Asp Asp Ala Ala Ile Ser Leu Ala Phe Ser<br>640 645 650     | 2088 |
| aaa aaa cag ata gat gat cga aag gaa tgg tta act aat ttc atg gag<br>Lys Lys Gln Ile Asp Asp Arg Lys Glu Trp Leu Thr Asn Phe Met Glu<br>655 660 665 670 | 2136 |
| gat aga aga caa cga aag tta ctt ggg ctt cct gag gat tac ttg tat<br>Asp Arg Arg Gln Arg Lys Leu Leu Gly Leu Pro Glu Asp Tyr Leu Tyr<br>675 680 685     | 2184 |
| gga caa act acc aca tat ctg aca tat aat gac ttc atc aac aag gaa<br>Gly Gln Thr Thr Tyr Leu Thr Tyr Asn Asp Phe Ile Asn Lys Glu<br>690 695 700         | 2232 |
| ctt atc ttg ttc tca aat tct gat aac gag aga tct atc cct tct atg<br>Leu Ile Leu Phe Ser Asn Ser Asp Asn Glu Arg Ser Ile Pro Ser Met<br>705 710 715     | 2280 |
| gtg gat ggt ttg aaa cca ggt cag aga aag gtt ttg ttt act tgc ttc<br>Val Asp Gly Leu Lys Pro Gly Gln Arg Lys Val Leu Phe Thr Cys Phe<br>720 725 730     | 2328 |
| aaa cgg aat gac aag cga gaa gta aag gtt gcc caa tta gct gga tca<br>Lys Arg Asn Asp Lys Arg Glu Val Lys Val Ala Gln Leu Ala Gly Ser<br>735 740 745 750 | 2376 |
| gtg gct gaa atg tct tct tat cat cat ggt gag atg tca cta atg atg<br>Val Ala Glu Met Ser Ser Tyr His His Gly Glu Met Ser Leu Met Met<br>755 760 765     | 2424 |
| acc att atc aat ttg gct cag aat ttt gtg ggt agc aat aat cta aac<br>Thr Ile Ile Asn Leu Ala Gln Asn Phe Val Gly Ser Asn Asn Leu Asn<br>770 775 780     | 2472 |
| ctc ttg cag ccc att ggt cag ttt ggt acc agg cta cat ggt ggc aag<br>Leu Leu Gln Pro Ile Gly Gln Phe Gly Thr Arg Leu His Gly Gly Lys<br>785 790 795     | 2520 |
| gat tct gct agt cca cga tac atc ttt aca atg ctc agc tct ttg gct<br>Asp Ser Ala Ser Pro Arg Tyr Ile Phe Thr Met Leu Ser Ser Leu Ala<br>800 805 810     | 2568 |
| cga ttg tta ttt cca cca aaa gat gat cac acg ttg aag ttt tta tat<br>Arg Leu Leu Phe Pro Pro Lys Asp Asp His Thr Leu Lys Phe Leu Tyr<br>815 820 825 830 | 2616 |
| gat gac aac cag cgt gtt gag cct gaa tgg tac att cct att att ccc<br>Asp Asp Asn Gln Arg Val Glu Pro Glu Trp Tyr Ile Pro Ile Ile Pro<br>835 840 845     | 2664 |
| atg gtg ctg ata aat ggt gct gaa gga atc ggt act ggg tgg tcc tgc<br>Met Val Leu Ile Asn Gly Ala Glu Gly Ile Gly Thr Gly Trp Ser Cys<br>850 855 860     | 2712 |

|   |      |
|---|------|
| aaa atc ccc aac ttt gat gtg cgt gaa att gta aat aac atc agg cgt<br>Lys Ile Pro Asn Phe Asp Val Arg Glu Ile Val Asn Asn Ile Arg Arg<br>865 870 875     | 2760 |
| ttg atg gat gga gaa gaa cct ttg cca atg ctt cca agt tac aag aac<br>Leu Met Asp Gly Glu Glu Pro Leu Pro Met Leu Pro Ser Tyr Lys Asn<br>880 885 890     | 2808 |
| ttc aag ggt act att gaa gaa ctg gct cca aat caa tat gtg att agt<br>Phe Lys Gly Thr Ile Glu Glu Leu Ala Pro Asn Gln Tyr Val Ile Ser<br>895 900 905 910 | 2856 |
| ggt gaa gta gct att ctt aat tct aca acc att gaa atc tca gag ctt<br>Gly Glu Val Ala Ile Leu Asn Ser Thr Thr Ile Glu Ile Ser Glu Leu<br>915 920 925     | 2904 |
| ccc gtc aga aca tgg acc cag aca tac aaa gaa caa gtt cta gaa ccc<br>Pro Val Arg Thr Trp Thr Gln Thr Tyr Lys Glu Gln Val Leu Glu Pro<br>930 935 940     | 2952 |
| atg ttg aat ggc acc gag aag aca cct cct ctc ata aca gac tat agg<br>Met Leu Asn Gly Thr Glu Lys Thr Pro Pro Leu Ile Thr Asp Tyr Arg<br>945 950 955     | 3000 |
| gaa tac cat aca gat acc act gtg aaa ttt gtt gtg aag atg act gaa<br>Glu Tyr His Thr Asp Thr Thr Val Lys Phe Val Val Lys Met Thr Glu<br>960 965 970     | 3048 |
| gaa aaa ctg gca gag gca gag aga gtt gga cta cac aaa gtc ttc aaa<br>Glu Lys Leu Ala Glu Ala Glu Arg Val Gly Leu His Lys Val Phe Lys<br>975 980 985 990 | 3096 |
| ctc caa act agt ctc aca tgc aac tct atg gtg ctt ttt gac cac gta<br>Leu Gln Thr Ser Leu Thr Cys Asn Ser Met Val Leu Phe Asp His Val<br>995 1000 1005   | 3144 |
| ggc tgt tta aag aaa tat gac acg gtg ttg gat att cta aga gac<br>Gly Cys Leu Lys Lys Tyr Asp Thr Val Leu Asp Ile Leu Arg Asp<br>1010 1015 1020          | 3189 |
| ttt ttt gaa ctc aga ctt aaa tat tat gga tta aga aaa gaa tgg<br>Phe Phe Glu Leu Arg Leu Lys Tyr Tyr Gly Leu Arg Lys Glu Trp<br>1025 1030 1035          | 3234 |
| ctc cta gga atg ctt ggt gct gaa tct gct aaa ctg aat aat cag<br>Leu Leu Gly Met Leu Gly Ala Glu Ser Ala Lys Leu Asn Asn Gln<br>1040 1045 1050          | 3279 |
| gct cgc ttt atc tta gag aaa ata gat ggc aaa ata atc att gaa<br>Ala Arg Phe Ile Leu Glu Lys Ile Asp Gly Lys Ile Ile Ile Glu<br>1055 1060 1065          | 3324 |
| aat aag cct aag aaa gaa tta att aaa gtt ctg att cag agg gga<br>Asn Lys Pro Lys Lys Glu Leu Ile Lys Val Leu Ile Gln Arg Gly<br>1070 1075 1080          | 3369 |
| tat gat tcg gat cct gtg aag gcc tgg aaa gaa gcc cag caa aag<br>Tyr Asp Ser Asp Pro Val Lys Ala Trp Lys Glu Ala Gln Gln Lys<br>1085 1090 1095          | 3414 |
| gtt cca gat gaa gaa gaa aat gaa gag agt gac aac gaa aag gaa<br>Val Pro Asp Glu Glu Glu Asn Glu Glu Ser Asp Asn Glu Lys Glu<br>1100 1105 1110          | 3459 |
| act gaa aag agt gac tcc gta aca gat tct gga cca acc ttc aac   | 3504 |

|                 |                     |                             |      |
|-----------------|---------------------|-----------------------------|------|
| Thr Glu Lys Ser | Asp Ser Val Thr     | Asp Ser Gly Pro Thr Phe Asn |      |
| 1115            | 1120                | 1125                        |      |
| tat ctt ctt gat | atg ccc ctt tgg tat | tta acc aag gaa aag aaa     | 3549 |
| Tyr Leu Leu Asp | Met Pro Leu Trp Tyr | Leu Thr Lys Glu Lys Lys     |      |
| 1130            | 1135                | 1140                        |      |
| gat gaa ctc tgc | agg cta aga aat gaa | aaa gaa caa gag ctg gac     | 3594 |
| Asp Glu Leu Cys | Arg Leu Arg Asn Glu | Lys Glu Gln Glu Leu Asp     |      |
| 1145            | 1150                | 1155                        |      |
| aca tta aaa aga | aag agt cca tca gat | ttg tgg aaa gaa gac ttg     | 3639 |
| Thr Leu Lys Arg | Lys Ser Pro Ser Asp | Leu Trp Lys Glu Asp Leu     |      |
| 1160            | 1165                | 1170                        |      |
| gct aca ttt att | gaa gaa ttg gag gct | gtt gaa gcc aag gaa aaa     | 3684 |
| Ala Thr Phe Ile | Glu Glu Leu Glu Ala | Val Glu Ala Lys Glu Lys     |      |
| 1175            | 1180                | 1185                        |      |
| caa gat gaa caa | gtc gga ctt cct ggg | aaa ggg ggg aag gcc aag     | 3729 |
| Gln Asp Glu Gln | Val Gly Leu Pro Gly | Lys Gly Gly Lys Ala Lys     |      |
| 1190            | 1195                | 1200                        |      |
| ggg aaa aaa aca | caa atg gct gaa gtt | ttg cct tct ccg cgt ggt     | 3774 |
| Gly Lys Lys Thr | Gln Met Ala Glu Val | Leu Pro Ser Pro Arg Gly     |      |
| 1205            | 1210                | 1215                        |      |
| caa aga gtc att | cca cga ata acc ata | gaa atg aaa gca gag gca     | 3819 |
| Gln Arg Val Ile | Pro Arg Ile Thr Ile | Glu Met Lys Ala Glu Ala     |      |
| 1220            | 1225                | 1230                        |      |
| gaa aag aaa aat | aaa aag aaa att aag | aat gaa aat act gaa gga     | 3864 |
| Glu Lys Lys Asn | Lys Lys Lys Ile Lys | Asn Glu Asn Thr Glu Gly     |      |
| 1235            | 1240                | 1245                        |      |
| agc cct caa gaa | gat ggt gtg gaa cta | gaa ggc cta aaa caa aga     | 3909 |
| Ser Pro Gln Glu | Asp Gly Val Glu Leu | Glu Gly Leu Lys Gln Arg     |      |
| 1250            | 1255                | 1260                        |      |
| tta gaa aag aaa | cag aaa aga gaa cca | ggt aca aag aca aag aaa     | 3954 |
| Leu Glu Lys Lys | Gln Lys Arg Glu Pro | Gly Thr Lys Thr Lys Lys     |      |
| 1265            | 1270                | 1275                        |      |
| caa act aca ttg | gca ttt aag cca atc | aaa aaa gga aag aag aga     | 3999 |
| Gln Thr Thr Leu | Ala Phe Lys Pro Ile | Lys Lys Gly Lys Lys Arg     |      |
| 1280            | 1285                | 1290                        |      |
| aat ccc tgg tct | gat tca gaa tca gat | agg agc agt gac gaa agt     | 4044 |
| Asn Pro Trp Ser | Asp Ser Glu Ser Asp | Arg Ser Ser Asp Glu Ser     |      |
| 1295            | 1300                | 1305                        |      |
| aat ttt gat gtc | cct cca cga gaa aca | gag cca cgg aga gca gca     | 4089 |
| Asn Phe Asp Val | Pro Pro Arg Glu Thr | Glu Pro Arg Arg Ala Ala     |      |
| 1310            | 1315                | 1320                        |      |
| aca aaa aca aaa | ttc aca atg gat ttg | gat tca gat gaa gat ttc     | 4134 |
| Thr Lys Thr Lys | Phe Thr Met Asp Leu | Asp Ser Asp Glu Asp Phe     |      |
| 1325            | 1330                | 1335                        |      |
| tca gat ttt gat | gaa aaa act gat gat | gaa gat ttt gtc cca tca     | 4179 |
| Ser Asp Phe Asp | Glu Lys Thr Asp Asp | Glu Asp Phe Val Pro Ser     |      |
| 1340            | 1345                | 1350                        |      |
| gat gct agt cca | cct aag acc aaa act | tcc cca aaa ctt agt aac     | 4224 |
| Asp Ala Ser Pro | Pro Lys Thr Lys Thr | Ser Pro Lys Leu Ser Asn     |      |
| 1355            | 1360                | 1365                        |      |

|  |      |
|--|------|
| aaa gaa ctg aaa cca cag aaa agt gtc gtg tca gac ctt gaa gct        | 4269 |
| Lys Glu Leu Lys Pro Gln Lys Ser Val Val Ser Asp Leu Glu Ala        |      |
| 1370 1375 1380   |      |
| gat gat gtt aag ggc agt gta cca ctg tct tca agc cct cct gct        | 4314 |
| Asp Asp Val Lys Gly Ser Val Pro Leu Ser Ser Ser Pro Pro Ala        |      |
| 1385 1390 1395   |      |
| aca cat ttc cca gat gaa act gaa att aca aac cca gtt cct aaa        | 4359 |
| Thr His Phe Pro Asp Glu Thr Glu Ile Thr Asn Pro Val Pro Lys        |      |
| 1400 1405 1410   |      |
| aag aat gtg aca gtg aag aag aca gca gca aaa agt cag tct tcc        | 4404 |
| Lys Asn Val Thr Val Lys Lys Thr Ala Ala Lys Ser Gln Ser Ser        |      |
| 1415 1420 1425   |      |
| acc tcc act acc ggt gcc aaa aaa agg gct gcc cca aaa gga act        | 4449 |
| Thr Ser Thr Thr Gly Ala Lys Lys Arg Ala Ala Pro Lys Gly Thr        |      |
| 1430 1435 1440   |      |
| aaa agg gat cca gct ttg aat tct ggt gtc tct caa aag cct gat        | 4494 |
| Lys Arg Asp Pro Ala Leu Asn Ser Gly Val Ser Gln Lys Pro Asp        |      |
| 1445 1450 1455   |      |
| cct gcc aaa acc aag aat cgc cgc aaa agg aag cca tcc act tot        | 4539 |
| Pro Ala Lys Thr Lys Asn Arg Arg Lys Arg Lys Pro Ser Thr Ser        |      |
| 1460 1465 1470   |      |
| gat gat tct gac tct aat ttt gag aaa att gtt tcg aaa gca gtc        | 4584 |
| Asp Asp Ser Asp Ser Asn Phe Glu Lys Ile Val Ser Lys Ala Val        |      |
| 1475 1480 1485   |      |
| aca agc aag aaa tcc aag ggg gag agt gat gac ttc cat atg gac        | 4629 |
| Thr Ser Lys Lys Ser Lys Gly Glu Ser Asp Asp Phe His Met Asp        |      |
| 1490 1495 1500   |      |
| ttt gac tca gct gtg gct cct cgg gca aaa tct gta cgg gca aag        | 4674 |
| Phe Asp Ser Ala Val Ala Pro Arg Ala Lys Ser Val Arg Ala Lys        |      |
| 1505 1510 1515   |      |
| aaa cct ata aag tac ctg gaa gag tca gat gaa gat gat ctg ttt        | 4719 |
| Lys Pro Ile Lys Tyr Leu Glu Glu Ser Asp Glu Asp Asp Leu Phe        |      |
| 1520 1525 1530   |      |
| taaaatgtga ggcgattatt ttaagtaatt atcttaccaa gcccaagact ggttttaaag  | 4779 |
| ttacctgaag ctcttaactt cctccctctt gaatttagtt tggggaaggt gtttttagta  | 4839 |
| caagacatca aagtgaagta aagccaagt gttcttttagc tttttataat actgtotaaa  | 4899 |
| tagtgaccat ctcattggga ttgttttctt ctctgctttg tctgtgtttt gagtctgctt  | 4959 |
| tcttttgtct ttaaaacctg atttttaagt tcttctgaac tgtagaaata gctatctgat  | 5019 |
| cacttcagcg taaagcagtg tgtttattaa ccatccacta agctaaaact agagcagttt  | 5079 |
| gatttaaaag tgtcactctt cctccttttc tactttcagt agatatgaga tagagcataa  | 5139 |
| ttatctgttt tatcttagtt ttatacataa ttaccatca gatagaactt tatggttcta   | 5199 |
| gtacagatac totactacac tcagcctctt atgtgccaaag tttttcttta agcaatgaga | 5259 |
| aattgctcat gttcttcac tctcaaate atcagaggcc aaagaaaaac actttggctg    | 5319 |
| tgtctataac ttgacacagt caatagaatg aagaaaatta gagtagttat gtgattattt  | 5379 |



cagctcttga cctgtcccct ctggtgcct ctgagtctga atctcccaa gagagaaacc 5439  
 aatttctaag aggactggat tgcagaagac tcggggacaa catttgatcc aagatcttaa 5499  
 atgttatatt gataaccatg ctcagcaatg agctattaga ttcattttgg gaaatctcca 5559  
 taatttcaat ttgtaaactt tgtaagacc tgtctacatt gttatatgtg tgtgacttga 5619  
 gtaatgttat caacgttttt gtaaataatt actatgtttt tctattagct aaattccaac 5679  
 aattttgtac ttttaataaa 5698

<210> 30

<211> 1531

<212> PRT

<213> NM\_001067 TOP2A

<400> 30

Met Glu Val Ser Pro Leu Gln Pro Val Asn Glu Asn Met Gln Val Asn  
 1 5 10 15

Lys Ile Lys Lys Asn Glu Asp Ala Lys Lys Arg Leu Ser Val Glu Arg  
 20 25 30

Ile Tyr Gln Lys Lys Thr Gln Leu Glu His Ile Leu Leu Arg Pro Asp  
 35 40 45

Thr Tyr Ile Gly Ser Val Glu Leu Val Thr Gln Gln Met Trp Val Tyr  
 50 55 60

Asp Glu Asp Val Gly Ile Asn Tyr Arg Glu Val Thr Phe Val Pro Gly  
 65 70 75 80

Leu Tyr Lys Ile Phe Asp Glu Ile Leu Val Asn Ala Ala Asp Asn Lys  
 85 90 95

Gln Arg Asp Pro Lys Met Ser Cys Ile Arg Val Thr Ile Asp Pro Glu  
 100 105 110

Asn Asn Leu Ile Ser Ile Trp Asn Asn Gly Lys Gly Ile Pro Val Val  
 115 120 125

Glu His Lys Val Glu Lys Met Tyr Val Pro Ala Leu Ile Phe Gly Gln  
 130 135 140

Leu Leu Thr Ser Ser Asn Tyr Asp Asp Asp Glu Lys Lys Val Thr Gly  
 145 150 155 160

Gly Arg Asn Gly Tyr Gly Ala Lys Leu Cys Asn Ile Phe Ser Thr Lys  
 165 170 175

Phe Thr Val Glu Thr Ala Ser Arg Glu Tyr Lys Lys Met Phe Lys Gln  
180 185 190

Thr Trp Met Asp Asn Met Gly Arg Ala Gly Glu Met Glu Leu Lys Pro  
195 200 205

Phe Asn Gly Glu Asp Tyr Thr Cys Ile Thr Phe Gln Pro Asp Leu Ser  
210 215 220

Lys Phe Lys Met Gln Ser Leu Asp Lys Asp Ile Val Ala Leu Met Val  
225 230 235 240

Arg Arg Ala Tyr Asp Ile Ala Gly Ser Thr Lys Asp Val Lys Val Phe  
245 250 255

Leu Asn Gly Asn Lys Leu Pro Val Lys Gly Phe Arg Ser Tyr Val Asp  
260 265 270

Met Tyr Leu Lys Asp Lys Leu Asp Glu Thr Gly Asn Ser Leu Lys Val  
275 280 285

Ile His Glu Gln Val Asn His Arg Trp Glu Val Cys Leu Thr Met Ser  
290 295 300

Glu Lys Gly Phe Gln Gln Ile Ser Phe Val Asn Ser Ile Ala Thr Ser  
305 310 315 320

Lys Gly Gly Arg His Val Asp Tyr Val Ala Asp Gln Ile Val Thr Lys  
325 330 335

Leu Val Asp Val Val Lys Lys Lys Asn Lys Gly Gly Val Ala Val Lys  
340 345 350

Ala His Gln Val Lys Asn His Met Trp Ile Phe Val Asn Ala Leu Ile  
355 360 365

Glu Asn Pro Thr Phe Asp Ser Gln Thr Lys Glu Asn Met Thr Leu Gln  
370 375 380

Pro Lys Ser Phe Gly Ser Thr Cys Gln Leu Ser Glu Lys Phe Ile Lys  
385 390 395 400

Ala Ala Ile Gly Cys Gly Ile Val Glu Ser Ile Leu Asn Trp Val Lys  
405 410 415

Phe Lys Ala Gln Val Gln Leu Asn Lys Lys Cys Ser Ala Val Lys His  
420 425 430

Asn Arg Ile Lys Gly Ile Pro Lys Leu Asp Asp Ala Asn Asp Ala Gly  
435 440 445

Gly Arg Asn Ser Thr Glu Cys Thr Leu Ile Leu Thr Glu Gly Asp Ser  
450 455 460

Ala Lys Thr Leu Ala Val Ser Gly Leu Gly Val Val Gly Arg Asp Lys  
465 470 475 480

Tyr Gly Val Phe Pro Leu Arg Gly Lys Ile Leu Asn Val Arg Glu Ala  
485 490 495

Ser His Lys Gln Ile Met Glu Asn Ala Glu Ile Asn Asn Ile Ile Lys  
500 505 510

Ile Val Gly Leu Gln Tyr Lys Lys Asn Tyr Glu Asp Glu Asp Ser Leu  
515 520 525

Lys Thr Leu Arg Tyr Gly Lys Ile Met Ile Met Thr Asp Gln Asp Gln  
530 535 540

Asp Gly Ser His Ile Lys Gly Leu Leu Ile Asn Phe Ile His His Asn  
545 550 555 560

Trp Pro Ser Leu Leu Arg His Arg Phe Leu Glu Glu Phe Ile Thr Pro  
565 570 575

Ile Val Lys Val Ser Lys Asn Lys Gln Glu Met Ala Phe Tyr Ser Leu  
580 585 590

Pro Glu Phe Glu Glu Trp Lys Ser Ser Thr Pro Asn His Lys Lys Trp  
595 600 605

Lys Val Lys Tyr Tyr Lys Gly Leu Gly Thr Ser Thr Ser Lys Glu Ala  
610 615 620

Lys Glu Tyr Phe Ala Asp Met Lys Arg His Arg Ile Gln Phe Lys Tyr  
625 630 635 640

Ser Gly Pro Glu Asp Asp Ala Ala Ile Ser Leu Ala Phe Ser Lys Lys  
645 650 655

Gln Ile Asp Asp Arg Lys Glu Trp Leu Thr Asn Phe Met Glu Asp Arg  
660 665 670

Arg Gln Arg Lys Leu Leu Gly Leu Pro Glu Asp Tyr Leu Tyr Gly Gln  
675 680 685

Thr Thr Thr Tyr Leu Thr Tyr Asn Asp Phe Ile Asn Lys Glu Leu Ile  
690 695 700

Leu Phe Ser Asn Ser Asp Asn Glu Arg Ser Ile Pro Ser Met Val Asp

705                      710                      715                      720

Gly Leu Lys Pro Gly Gln Arg Lys Val Leu Phe Thr Cys Phe Lys Arg  
725                      730                      735

Asn Asp Lys Arg Glu Val Lys Val Ala Gln Leu Ala Gly Ser Val Ala  
740                      745                      750

Glu Met Ser Ser Tyr His His Gly Glu Met Ser Leu Met Met Thr Ile  
755                      760                      765

Ile Asn Leu Ala Gln Asn Phe Val Gly Ser Asn Asn Leu Asn Leu Leu  
770                      775                      780

Gln Pro Ile Gly Gln Phe Gly Thr Arg Leu His Gly Gly Lys Asp Ser  
785                      790                      795                      800

Ala Ser Pro Arg Tyr Ile Phe Thr Met Leu Ser Ser Leu Ala Arg Leu  
805                      810                      815

Leu Phe Pro Pro Lys Asp Asp His Thr Leu Lys Phe Leu Tyr Asp Asp  
820                      825                      830

Asn Gln Arg Val Glu Pro Glu Trp Tyr Ile Pro Ile Ile Pro Met Val  
835                      840                      845

Leu Ile Asn Gly Ala Glu Gly Ile Gly Thr Gly Trp Ser Cys Lys Ile  
850                      855                      860

Pro Asn Phe Asp Val Arg Glu Ile Val Asn Asn Ile Arg Arg Leu Met  
865                      870                      875                      880

Asp Gly Glu Glu Pro Leu Pro Met Leu Pro Ser Tyr Lys Asn Phe Lys  
885                      890                      895

Gly Thr Ile Glu Glu Leu Ala Pro Asn Gln Tyr Val Ile Ser Gly Glu  
900                      905                      910

Val Ala Ile Leu Asn Ser Thr Thr Ile Glu Ile Ser Glu Leu Pro Val  
915                      920                      925

Arg Thr Trp Thr Gln Thr Tyr Lys Glu Gln Val Leu Glu Pro Met Leu  
930                      935                      940

Asn Gly Thr Glu Lys Thr Pro Pro Leu Ile Thr Asp Tyr Arg Glu Tyr  
945                      950                      955                      960

His Thr Asp Thr Thr Val Lys Phe Val Val Lys Met Thr Glu Glu Lys  
965                      970                      975

Leu Ala Glu Ala Glu Arg Val Gly Leu His Lys Val Phe Lys Leu Gln

980                      985                      990  
 Thr Ser Leu Thr Cys Asn Ser Met Val Leu Phe Asp His Val Gly Cys  
           995                                      1000                                      1005  
 Leu Lys Lys Tyr Asp Thr Val Leu Asp Ile Leu Arg Asp Phe Phe  
           1010                                      1015                                      1020  
 Glu Leu Arg Leu Lys Tyr Tyr Gly Leu Arg Lys Glu Trp Leu Leu  
           1025                                      1030                                      1035  
 Gly Met Leu Gly Ala Glu Ser Ala Lys Leu Asn Asn Gln Ala Arg  
           1040                                      1045                                      1050  
 Phe Ile Leu Glu Lys Ile Asp Gly Lys Ile Ile Ile Glu Asn Lys  
           1055                                      1060                                      1065  
 Pro Lys Lys Glu Leu Ile Lys Val Leu Ile Gln Arg Gly Tyr Asp  
           1070                                      1075                                      1080  
 Ser Asp Pro Val Lys Ala Trp Lys Glu Ala Gln Gln Lys Val Pro  
           1085                                      1090                                      1095  
 Asp Glu Glu Glu Asn Glu Glu Ser Asp Asn Glu Lys Glu Thr Glu  
           1100                                      1105                                      1110  
 Lys Ser Asp Ser Val Thr Asp Ser Gly Pro Thr Phe Asn Tyr Leu  
           1115                                      1120                                      1125  
 Leu Asp Met Pro Leu Trp Tyr Leu Thr Lys Glu Lys Lys Asp Glu  
           1130                                      1135                                      1140  
 Leu Cys Arg Leu Arg Asn Glu Lys Glu Gln Glu Leu Asp Thr Leu  
           1145                                      1150                                      1155  
 Lys Arg Lys Ser Pro Ser Asp Leu Trp Lys Glu Asp Leu Ala Thr  
           1160                                      1165                                      1170  
 Phe Ile Glu Glu Leu Glu Ala Val Glu Ala Lys Glu Lys Gln Asp  
           1175                                      1180                                      1185  
 Glu Gln Val Gly Leu Pro Gly Lys Gly Gly Lys Ala Lys Gly Lys  
           1190                                      1195                                      1200  
 Lys Thr Gln Met Ala Glu Val Leu Pro Ser Pro Arg Gly Gln Arg  
           1205                                      1210                                      1215  
 Val Ile Pro Arg Ile Thr Ile Glu Met Lys Ala Glu Ala Glu Lys  
           1220                                      1225                                      1230  
 Lys Asn Lys Lys Lys Ile Lys Asn Glu Asn Thr Glu Gly Ser Pro

1235                      1240                      1245

Gln Glu Asp Gly Val Glu Leu Glu Gly Leu Lys Gln Arg Leu Glu  
1250                      1255                      1260

Lys Lys Gln Lys Arg Glu Pro Gly Thr Lys Thr Lys Lys Gln Thr  
1265                      1270                      1275

Thr Leu Ala Phe Lys Pro Ile Lys Lys Gly Lys Lys Arg Asn Pro  
1280                      1285                      1290

Trp Ser Asp Ser Glu Ser Asp Arg Ser Ser Asp Glu Ser Asn Phe  
1295                      1300                      1305

Asp Val Pro Pro Arg Glu Thr Glu Pro Arg Arg Ala Ala Thr Lys  
1310                      1315                      1320

Thr Lys Phe Thr Met Asp Leu Asp Ser Asp Glu Asp Phe Ser Asp  
1325                      1330                      1335

Phe Asp Glu Lys Thr Asp Asp Glu Asp Phe Val Pro Ser Asp Ala  
1340                      1345                      1350

Ser Pro Pro Lys Thr Lys Thr Ser Pro Lys Leu Ser Asn Lys Glu  
1355                      1360                      1365

Leu Lys Pro Gln Lys Ser Val Val Ser Asp Leu Glu Ala Asp Asp  
1370                      1375                      1380

Val Lys Gly Ser Val Pro Leu Ser Ser Ser Pro Pro Ala Thr His  
1385                      1390                      1395

Phe Pro Asp Glu Thr Glu Ile Thr Asn Pro Val Pro Lys Lys Asn  
1400                      1405                      1410

Val Thr Val Lys Lys Thr Ala Ala Lys Ser Gln Ser Ser Thr Ser  
1415                      1420                      1425

Thr Thr Gly Ala Lys Lys Arg Ala Ala Pro Lys Gly Thr Lys Arg  
1430                      1435                      1440

Asp Pro Ala Leu Asn Ser Gly Val Ser Gln Lys Pro Asp Pro Ala  
1445                      1450                      1455

Lys Thr Lys Asn Arg Arg Lys Arg Lys Pro Ser Thr Ser Asp Asp  
1460                      1465                      1470

Ser Asp Ser Asn Phe Glu Lys Ile Val Ser Lys Ala Val Thr Ser  
1475                      1480                      1485

Lys Lys Ser Lys Gly Glu Ser Asp Asp Phe His Met Asp Phe Asp  
1490 1495 1500

Ser Ala Val Ala Pro Arg Ala Lys Ser Val Arg Ala Lys Lys Pro  
1505 1510 1515

Ile Lys Tyr Leu Glu Glu Ser Asp Glu Asp Asp Leu Phe  
1520 1525 1530

<210> 31

<211> 1369

<212> DNA

<213> NM\_138410 chemokine-like factor super family 7, CKLFSF7

<220>

<221> CDS

<222> (237)..(761)

<223>

<400> 31

agtctaactt ccgcgcattt acgagggccg ggactgcccgc taccttttctg gaaggcgccg 60

gccggccagt caccggaggg atcccgccag aggcggctcc gcgtctccag cagcggggca 120

gggacccggg cgccccgcc tcgccagcgc ccgccccgc ccgccccgg cccgccctct 180

gtatctggcc cctgggcagc tgccgggga ggcggccagc gagctggggc cgcgca atg 239  
Met  
1

tcg cac gga gcc ggg ctc gtc cgc acc acg tgc agc agc ggc agc gcg 287  
Ser His Gly Ala Gly Leu Val Arg Thr Thr Cys Ser Ser Gly Ser Ala  
5 10 15

ctc gga ccc ggg gcc ggc gcg gcc cag ccc agc gcg agc ccc ttg gag 335  
Leu Gly Pro Gly Ala Gly Ala Gln Pro Ser Ala Ser Pro Leu Glu  
20 25 30

ggg ctg ctg gac ctc agc tac ccc cgc acc cac gcg gcc ctg ctg aaa 383  
Gly Leu Leu Asp Leu Ser Tyr Pro Arg Thr His Ala Ala Leu Leu Lys  
35 40 45

gtg gcg caa atg gtc acc ctg ctg att gcc ttc atc tgt gtg cgg agc 431  
Val Ala Gln Met Val Thr Leu Leu Ile Ala Phe Ile Cys Val Arg Ser  
50 55 60 65

tcc ctg tgg acc aac tac agc gcc tac agc tac ttt gaa gtg gtc acc 479  
Ser Leu Trp Thr Asn Tyr Ser Ala Tyr Ser Tyr Phe Glu Val Val Thr  
70 75 80

att tgc gac ttg ata atg atc ctc gcc ttt tac ctg gtc cac ctc ttc 527  
Ile Cys Asp Leu Ile Met Ile Leu Ala Phe Tyr Leu Val His Leu Phe  
85 90 95

cgc ttc tac cgc gtg ctc acc tgt atc agc tgg ccc ctg tcg gaa ctt 575

Arg Phe Tyr Arg Val Leu Thr Cys Ile Ser Trp Pro Leu Ser Glu Leu  
 100 105 110  
 ctg cac tat tta atc ggt acc ctg ctc ctc ctc atc gcc tcc att gtg 623  
 Leu His Tyr Leu Ile Gly Thr Leu Leu Leu Ile Ala Ser Ile Val  
 115 120 125  
 gca gct tcc aag agt tac aac cag agc gga ctg gta gcc gga gcg atc 671  
 Ala Ala Ser Lys Ser Tyr Asn Gln Ser Gly Leu Val Ala Gly Ala Ile  
 130 135 140 145  
 ttt ggt ttc atg gcc acc ttc ctc tgc atg gca agc ata tgg ctg tcc 719  
 Phe Gly Phe Met Ala Thr Phe Leu Cys Met Ala Ser Ile Trp Leu Ser  
 150 155 160  
 tat aag atc tcg tgt gta acc cag tcc aca gat gca gcc gtc 761  
 Tyr Lys Ile Ser Cys Val Thr Gln Ser Thr Asp Ala Ala Val  
 165 170 175  
 tgatgaggcc acaaccctca ggcctctcag gagctttgca gagaggagga cgtgtactcc 821  
 aggcgaggcc tctggacctg tggtctctgt ccaaagtcct gtcaggctgg tgggcaccag 881  
 gaaaggctg caccctcttc ctgctctccc aggaagccag ctccctgagc tcctgagcca 941  
 gccggaaact ctctctccag ccttccgggg agaacatccc tccattctg ggaaaggaaa 1001  
 gcagctccca gggaaatgtt ttctgccttc ctgcttctag aaccacotca ggtactgatg 1061  
 aaccccaact agcacagctg aaggggtttg tgaatactcc cgcctaaatc ctttctactt 1121  
 cactctctcag gggagtgaag tgccttaaga aacaaagccc tgccttaatt tatctagctt 1181  
 gtcagtcagg tcttagagat accctctttc ctgaagttag gcgtgcctgt agaaacacta 1241  
 tgtggtcagc ctgtccccaa ggagatcttg tgtctcctct ccattctctgc ctttgttacc 1301  
 agtgtgcatg tgtttgtgtg ttttttaata aaatattgac tcggccagtt aaaaaaaaaa 1361  
 aaaaaaaaaa 1369  
 <210> 32  
 <211> 175  
 <212> PRT  
 <213> NM\_138410 chemokine-like factor super family 7, CKLFSP7  
 <400> 32  
 Met Ser His Gly Ala Gly Leu Val Arg Thr Thr Cys Ser Ser Gly Ser  
 1 5 10 15  
 Ala Leu Gly Pro Gly Ala Gly Ala Ala Gln Pro Ser Ala Ser Pro Leu  
 20 25 30  
 Glu Gly Leu Leu Asp Leu Ser Tyr Pro Arg Thr His Ala Ala Leu Leu  
 35 40 45  
 Lys Val Ala Gln Met Val Thr Leu Leu Ile Ala Phe Ile Cys Val Arg  
 50 55 60



Ser Ser Leu Trp Thr Asn Tyr Ser Ala Tyr Ser Tyr Phe Glu Val Val  
65 70 75 80

Thr Ile Cys Asp Leu Ile Met Ile Leu Ala Phe Tyr Leu Val His Leu  
85 90 95

Phe Arg Phe Tyr Arg Val Leu Thr Cys Ile Ser Trp Pro Leu Ser Glu  
100 105 110

Leu Leu His Tyr Leu Ile Gly Thr Leu Leu Leu Leu Ile Ala Ser Ile  
115 120 125

Val Ala Ala Ser Lys Ser Tyr Asn Gln Ser Gly Leu Val Ala Gly Ala  
130 135 140

Ile Phe Gly Phe Met Ala Thr Phe Leu Cys Met Ala Ser Ile Trp Leu  
145 150 155 160

Ser Tyr Lys Ile Ser Cys Val Thr Gln Ser Thr Asp Ala Ala Val  
165 170 175

<210> 33

<211> 3470

<212> DNA

<213> NM\_012112 C20orf1, chromosome 20 open reading frame 1

<220>

<221> CDS

<222> (458) .. (2698)

<223>

<400> 33

```

ggctaggtga gccgtgggaa gaaaagaggg agcagctagg gcgcgggtct ccctcctccc      60
ggagtttggg acggctgaag ttcaccttcc agcccctagc gccgttcgcg ccgctaggcc      120
tggcttctga ggcggttgcg gtgctcggtc gccgcctagg cggggcaggg tgcgagcagg      180
ggcttcgggc cacgcttctc ttggcgacag gattttgctg tgaagtccgt ccgggaaacg      240
gaggaaaaaa agagttgcgg gaggtgtcgc gctaataacg gttcttgata catatttgcc      300
agacttcaag atttcagaaa aggggtgaaa gagaagattg caactttgag tcagacctgt      360
aggcctgata gactgattaa accacagaag gtgacctgct gagaaaagtg gtacaaatac      420
tgggaaaaaac ctgctcttct gcgttaagtg ggagaca atg tca caa gtt aaa agc      475
                               Met Ser Gln Val Lys Ser
                               1           5

```

|   |      |
|---|------|
| tct tat tcc tat gat gcc ccc tcg gat ttc atc aat ttt tca tcc ttg | 523  |
| Ser Tyr Ser Tyr Asp Ala Pro Ser Asp Phe Ile Asn Phe Ser Ser Leu |      |
| 10 15 20  |      |
| gat gat gaa gga gat act caa aac ata gat tca tgg ttt gag gag aag | 571  |
| Asp Asp Glu Gly Asp Thr Gln Asn Ile Asp Ser Trp Phe Glu Glu Lys |      |
| 25 30 35  |      |
| gcc aat ttg gag aat aag tta ctg ggg aag aat gga act gga ggg ctt | 619  |
| Ala Asn Leu Glu Asn Lys Leu Leu Gly Lys Asn Gly Thr Gly Gly Leu |      |
| 40 45 50  |      |
| ttt cag ggc aaa act cct ttg aga aag gct aat ctt cag caa gct att | 667  |
| Phe Gln Gly Lys Thr Pro Leu Arg Lys Ala Asn Leu Gln Gln Ala Ile |      |
| 55 60 65 70   |      |
| gtc aca cct ttg aaa cca gtt gac aac act tac tac aaa gag gca gaa | 715  |
| Val Thr Pro Leu Lys Pro Val Asp Asn Thr Tyr Tyr Lys Glu Ala Glu |      |
| 75 80 85  |      |
| aaa gaa aat ctt gtg gaa caa tcc att ccg tca aat gct tgt tct tcc | 763  |
| Lys Glu Asn Leu Val Glu Gln Ser Ile Pro Ser Asn Ala Cys Ser Ser |      |
| 90 95 100   |      |
| ctg gaa gtt gag gca gcc ata tca aga aaa act cca gcc cag cct cag | 811  |
| Leu Glu Val Glu Ala Ala Ile Ser Arg Lys Thr Pro Ala Gln Pro Gln |      |
| 105 110 115   |      |
| aga aga tct ctt agg ctt tct gct cag aag gat ttg gaa cag aaa gaa | 859  |
| Arg Arg Ser Leu Arg Leu Ser Ala Gln Lys Asp Leu Glu Gln Lys Glu |      |
| 120 125 130   |      |
| aag cat cat gta aaa atg aaa gcc aag aga tgt gcc act cct gta atc | 907  |
| Lys His His Val Lys Met Lys Ala Lys Arg Cys Ala Thr Pro Val Ile |      |
| 135 140 145 150   |      |
| atc gat gaa att cta ccc tct aag aaa atg aaa gtt tct aac aac aaa | 955  |
| Ile Asp Glu Ile Leu Pro Ser Lys Lys Met Lys Val Ser Asn Asn Lys |      |
| 155 160 165   |      |
| aag aag cca gag gaa gaa ggc agt gct cat caa gat act gct gaa aag | 1003 |
| Lys Lys Pro Glu Glu Glu Gly Ser Ala His Gln Asp Thr Ala Glu Lys |      |
| 170 175 180   |      |
| aat gca tct tcc cca gag aaa gcc aag ggt aga cat act gtg cct tgt | 1051 |
| Asn Ala Ser Ser Pro Glu Lys Ala Lys Gly Arg His Thr Val Pro Cys |      |
| 185 190 195   |      |
| atg cca cct gca aag cag aag ttt cta aaa agt act gag gag caa gag | 1099 |
| Met Pro Pro Ala Lys Gln Lys Phe Leu Lys Ser Thr Glu Glu Gln Glu |      |
| 200 205 210   |      |
| ctg gag aag agt atg aaa atg cag caa gag gtg gtg gag atg cgg aaa | 1147 |
| Leu Glu Lys Ser Met Lys Met Gln Gln Glu Val Val Glu Met Arg Lys |      |
| 215 220 225 230   |      |
| aag aat gaa gaa ttc aag aaa ctt gct ctg gct gga ata ggg caa cct | 1195 |
| Lys Asn Glu Glu Phe Lys Lys Leu Ala Leu Ala Gly Ile Gly Gln Pro |      |
| 235 240 245   |      |
| gtg aag aaa tca gtg agc cag gtc acc aaa tca gtt gac ttc cac ttc | 1243 |
| Val Lys Lys Ser Val Ser Gln Val Thr Lys Ser Val Asp Phe His Phe |      |
| 250 255 260   |      |
| cgc aca gat gag cga atc aaa caa cat cct aag aac cag gag gaa tat | 1291 |
| Arg Thr Asp Glu Arg Ile Lys Gln His Pro Lys Asn Gln Glu Glu Tyr |      |
| 265 270 275   |      |

|   |      |
|---|------|
| aag gaa gtg aac ttt aca tct gaa cta cga aag cat cct tca tct cct<br>Lys Glu Val Asn Phe Thr Ser Glu Leu Arg Lys His Pro Ser Ser Pro<br>280 285 290     | 1339 |
| gcc cga gtg act aag gga tgt acc att gtt aag cct ttc aac ctg tcc<br>Ala Arg Val Thr Lys Gly Cys Thr Ile Val Lys Pro Phe Asn Leu Ser<br>295 300 305 310 | 1387 |
| caa gga aag aaa aga aca ttt gat gaa aca gtt tct aca tat gtg ccc<br>Gln Gly Lys Lys Arg Thr Phe Asp Glu Thr Val Ser Thr Tyr Val Pro<br>315 320 325     | 1435 |
| ctt gca cag caa gtt gaa gac ttc cat aaa cga acc cct aac aga tat<br>Leu Ala Gln Gln Val Glu Asp Phe His Lys Arg Thr Pro Asn Arg Tyr<br>330 335 340     | 1483 |
| cat ttg agg agc aag aag gat gat att aac ctg tta ccc tcc aaa tct<br>His Leu Arg Ser Lys Lys Asp Asp Ile Asn Leu Leu Pro Ser Lys Ser<br>345 350 355     | 1531 |
| tct gtg acc aag att tgc aga gac cca cag act cct gta ctg caa acc<br>Ser Val Thr Lys Ile Cys Arg Asp Pro Gln Thr Pro Val Leu Gln Thr<br>360 365 370     | 1579 |
| aaa cac cgt gca cgg gct gtg acc tgc aaa agt aca gca gag ctg gag<br>Lys His Arg Ala Arg Val Thr Cys Lys Ser Thr Ala Glu Leu Glu<br>375 380 385 390     | 1627 |
| gct gag gag ctc gag aaa ttg caa caa tac aaa ttc aaa gca cgt gaa<br>Ala Glu Glu Leu Glu Lys Leu Gln Gln Tyr Lys Phe Lys Ala Arg Glu<br>395 400 405     | 1675 |
| ctt gat ccc aga ata ctt gaa ggt ggg ccc atc ttg ccc aag aaa cca<br>Leu Asp Pro Arg Ile Leu Glu Gly Gly Pro Ile Leu Pro Lys Lys Pro<br>410 415 420     | 1723 |
| cct gtg aaa cca ccc acc gag cct att ggc ttt gat ttg gaa att gag<br>Pro Val Lys Pro Pro Thr Glu Pro Ile Gly Phe Asp Leu Glu Ile Glu<br>425 430 435     | 1771 |
| aaa aga atc cag gag cga gaa tca aag aag aaa aca gag gat gaa cac<br>Lys Arg Ile Gln Glu Arg Glu Ser Lys Lys Lys Thr Glu Asp Glu His<br>440 445 450     | 1819 |
| ttt gaa ttt cat tcc aga cct tgc cct act aag att ttg gaa gat gtt<br>Phe Glu Phe His Ser Arg Pro Cys Pro Thr Lys Ile Leu Glu Asp Val<br>455 460 465 470 | 1867 |
| gtg ggt gtt cct gaa aag aag gta ctt cca atc acc gtc ccc aag tca<br>Val Gly Val Pro Glu Lys Lys Val Leu Pro Ile Thr Val Pro Lys Ser<br>475 480 485     | 1915 |
| cca gcc ttt gca ttg aag aac aga att cga atg ccc acc aaa gaa gat<br>Pro Ala Phe Ala Leu Lys Asn Arg Ile Arg Met Pro Thr Lys Glu Asp<br>490 495 500     | 1963 |
| gag gaa gag gac gaa ccg gta gtg ata aaa gct caa cct gtg cca cat<br>Glu Glu Glu Asp Glu Pro Val Val Ile Lys Ala Gln Pro Val Pro His<br>505 510 515     | 2011 |
| tat ggg gtg cct ttt aag ccc caa atc cca gag gca aga act gtg gaa<br>Tyr Gly Val Pro Phe Lys Pro Gln Ile Pro Glu Ala Arg Thr Val Glu<br>520 525 530     | 2059 |
| ata tgc cct ttc tcg ttt gat tct cga gac aaa gaa cgt cag tta cag<br>Ile Cys Pro Phe Ser Phe Asp Ser Arg Asp Lys Glu Arg Gln Leu Gln<br>535 540 545 550 | 2107 |

|   |      |
|---|------|
| aag gag aag aaa ata aaa gaa ctg cag aaa ggg gag gtg ccc aag ttc<br>Lys Glu Lys Lys Ile Lys Glu Leu Gln Lys Gly Glu Val Pro Lys Phe<br>555 560 565     | 2155 |
| aag gca ctt ccc ttg cct cat ttt gac acc att aac ctg cca gag aag<br>Lys Ala Leu Pro Leu Pro His Phe Asp Thr Ile Asn Leu Pro Glu Lys<br>570 575 580     | 2203 |
| aag gta aag aat gtg acc cag att gaa cct ttc tgc ttg gag act gac<br>Lys Val Lys Asn Val Thr Gln Ile Glu Pro Phe Cys Leu Glu Thr Asp<br>585 590 595     | 2251 |
| aga aga ggt gct ctg aag gca cag act tgg aag cac cag ctg gaa gaa<br>Arg Arg Gly Ala Leu Lys Ala Gln Thr Trp Lys His Gln Leu Glu Glu<br>600 605 610     | 2299 |
| gaa ctg aga cag cag aaa gaa gca gct tgt ttc aag gct cgt cca aac<br>Glu Leu Arg Gln Gln Lys Glu Ala Ala Cys Phe Lys Ala Arg Pro Asn<br>615 620 625 630 | 2347 |
| acc gtc atc tct cag gag ccc ttt gtt ccc aag aaa gag aag aaa tca<br>Thr Val Ile Ser Gln Glu Pro Phe Val Pro Lys Lys Glu Lys Lys Ser<br>635 640 645     | 2395 |
| gtt gct gag ggc ctt tct ggt tct cta gtt cag gaa cct ttt cag ctg<br>Val Ala Glu Gly Leu Ser Gly Ser Leu Val Gln Glu Pro Phe Gln Leu<br>650 655 660     | 2443 |
| gct act gag aag aga gcc aaa gag cgg cag gag ctg gag aag aga atg<br>Ala Thr Glu Lys Arg Ala Lys Glu Arg Gln Glu Leu Glu Lys Arg Met<br>665 670 675     | 2491 |
| gct gag gta gaa gcc cag aaa gcc cag cag ttg gag gag gcc aga cta<br>Ala Glu Val Glu Ala Gln Lys Ala Gln Gln Leu Glu Glu Ala Arg Leu<br>680 685 690     | 2539 |
| cag gag gaa gag cag aaa aaa gag gag ctg gcc agg cta cgg aga gaa<br>Gln Glu Glu Glu Gln Lys Lys Glu Glu Leu Ala Arg Leu Arg Arg Glu<br>695 700 705 710 | 2587 |
| ctg gtg cat aag gca aat cca ata cgc aag tac cag ggt ctg gag ata<br>Leu Val His Lys Ala Asn Pro Ile Arg Lys Tyr Gln Gly Leu Glu Ile<br>715 720 725     | 2635 |
| aag tca agt gac cag cct ctg act gtg cct gta tct ccc aaa ttc tcc<br>Lys Ser Ser Asp Gln Pro Leu Thr Val Pro Val Ser Pro Lys Phe Ser<br>730 735 740     | 2683 |
| act cga ttc cac tgc taaactcagc tgtgagctgc ggataccgcc cggcaatggg<br>Thr Arg Phe His Cys<br>745   | 2738 |
| acctgctctt aacctcaaac ctaggaccgt cttgctttgt cattgggcat ggagagaacc   | 2798 |
| cattttctcca gacttttacc taccctgtgcc tgagaaagca tacttgacaa ctgtggactc   | 2858 |
| cagttttgtt gagaattgtt ttcttacatt actaaggcta ataatgagat gtaactcatg   | 2918 |
| aatgtctcga ttagactcca tgtagtact tcctttaaac catcagccgg ccttttatat  | 2978 |
| gggtcttcac tctgactaga atttagtctc tgtgtcagca cagtgtaatc tctattgcta   | 3038 |
| ttgccctta cgactctcac cctctcccca ctttttttaa aaattttaac cagaaaataa  | 3098 |
| agatagttaa atcctaagat agagattaag tcatgggtta aatgaggaa aatcagtaaa  | 3158 |
| tcagattctg tcctcttctc tgcataccgt gaatttatag ttaaggatcc ctttgctgtg   | 3218 |

```

agggtagaaa acctcaccaa ctgcaccagt gaggaagaag actgcgtgga ttcattgggga 3278
gcttcacagc agccacgcag caggctctgg gtggggctgc cgtaaggca cgttctttcc 3338
ttactggtgc tgataacaac agggaaccgt gcagtgtgca ttttaagacc tggcctggaa 3398
taaatacgtt ttgtctttcc ctcaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3458
aaaaaaaaaa aa 3470

```

&lt;210&gt; 34

&lt;211&gt; 747

&lt;212&gt; PRT

&lt;213&gt; NM\_012112 C20orf1, chromosome 20 open reading frame 1

&lt;400&gt; 34

```

Met Ser Gln Val Lys Ser Ser Tyr Ser Tyr Asp Ala Pro Ser Asp Phe
1             5             10             15

```

```

Ile Asn Phe Ser Ser Leu Asp Asp Glu Gly Asp Thr Gln Asn Ile Asp
                20             25             30

```

```

Ser Trp Phe Glu Glu Lys Ala Asn Leu Glu Asn Lys Leu Leu Gly Lys
            35             40             45

```

```

Asn Gly Thr Gly Gly Leu Phe Gln Gly Lys Thr Pro Leu Arg Lys Ala
50             55             60

```

```

Asn Leu Gln Gln Ala Ile Val Thr Pro Leu Lys Pro Val Asp Asn Thr
65             70             75             80

```

```

Tyr Tyr Lys Glu Ala Glu Lys Glu Asn Leu Val Glu Gln Ser Ile Pro
            85             90             95

```

```

Ser Asn Ala Cys Ser Ser Leu Glu Val Glu Ala Ala Ile Ser Arg Lys
100             105             110

```

```

Thr Pro Ala Gln Pro Gln Arg Arg Ser Leu Arg Leu Ser Ala Gln Lys
115             120             125

```

```

Asp Leu Glu Gln Lys Glu Lys His His Val Lys Met Lys Ala Lys Arg
130             135             140

```

```

Cys Ala Thr Pro Val Ile Ile Asp Glu Ile Leu Pro Ser Lys Lys Met
145             150             155             160

```

```

Lys Val Ser Asn Asn Lys Lys Lys Pro Glu Glu Glu Gly Ser Ala His
165             170             175

```

Gln Asp Thr Ala Glu Lys Asn Ala Ser Ser Pro Glu Lys Ala Lys Gly  
180 185 190

Arg His Thr Val Pro Cys Met Pro Pro Ala Lys Gln Lys Phe Leu Lys  
195 200 205

Ser Thr Glu Glu Gln Glu Leu Glu Lys Ser Met Lys Met Gln Gln Glu  
210 215 220

Val Val Glu Met Arg Lys Lys Asn Glu Glu Phe Lys Lys Leu Ala Leu  
225 230 235 240

Ala Gly Ile Gly Gln Pro Val Lys Lys Ser Val Ser Gln Val Thr Lys  
245 250 255

Ser Val Asp Phe His Phe Arg Thr Asp Glu Arg Ile Lys Gln His Pro  
260 265 270

Lys Asn Gln Glu Glu Tyr Lys Glu Val Asn Phe Thr Ser Glu Leu Arg  
275 280 285

Lys His Pro Ser Ser Pro Ala Arg Val Thr Lys Gly Cys Thr Ile Val  
290 295 300

Lys Pro Phe Asn Leu Ser Gln Gly Lys Lys Arg Thr Phe Asp Glu Thr  
305 310 315 320

Val Ser Thr Tyr Val Pro Leu Ala Gln Gln Val Glu Asp Phe His Lys  
325 330 335

Arg Thr Pro Asn Arg Tyr His Leu Arg Ser Lys Lys Asp Asp Ile Asn  
340 345 350

Leu Leu Pro Ser Lys Ser Ser Val Thr Lys Ile Cys Arg Asp Pro Gln  
355 360 365

Thr Pro Val Leu Gln Thr Lys His Arg Ala Arg Ala Val Thr Cys Lys  
370 375 380

Ser Thr Ala Glu Leu Glu Ala Glu Glu Leu Glu Lys Leu Gln Gln Tyr  
385 390 395 400

Lys Phe Lys Ala Arg Glu Leu Asp Pro Arg Ile Leu Glu Gly Gly Pro  
405 410 415

Ile Leu Pro Lys Lys Pro Pro Val Lys Pro Pro Thr Glu Pro Ile Gly  
420 425 430

Phe Asp Leu Glu Ile Glu Lys Arg Ile Gln Glu Arg Glu Ser Lys Lys  
435 440 445

Lys Thr Glu Asp Glu His Phe Glu Phe His Ser Arg Pro Cys Pro Thr  
450 455 460

Lys Ile Leu Glu Asp Val Val Gly Val Pro Glu Lys Lys Val Leu Pro  
465 470 475 480

Ile Thr Val Pro Lys Ser Pro Ala Phe Ala Leu Lys Asn Arg Ile Arg  
485 490 495

Met Pro Thr Lys Glu Asp Glu Glu Glu Asp Glu Pro Val Val Ile Lys  
500 505 510

Ala Gln Pro Val Pro His Tyr Gly Val Pro Phe Lys Pro Gln Ile Pro  
515 520 525

Glu Ala Arg Thr Val Glu Ile Cys Pro Phe Ser Phe Asp Ser Arg Asp  
530 535 540

Lys Glu Arg Gln Leu Gln Lys Glu Lys Lys Ile Lys Glu Leu Gln Lys  
545 550 555 560

Gly Glu Val Pro Lys Phe Lys Ala Leu Pro Leu Pro His Phe Asp Thr  
565 570 575

Ile Asn Leu Pro Glu Lys Lys Val Lys Asn Val Thr Gln Ile Glu Pro  
580 585 590

Phe Cys Leu Glu Thr Asp Arg Arg Gly Ala Leu Lys Ala Gln Thr Trp  
595 600 605

Lys His Gln Leu Glu Glu Glu Leu Arg Gln Gln Lys Glu Ala Ala Cys  
610 615 620

Phe Lys Ala Arg Pro Asn Thr Val Ile Ser Gln Glu Pro Phe Val Pro  
625 630 635 640

Lys Lys Glu Lys Lys Ser Val Ala Glu Gly Leu Ser Gly Ser Leu Val  
645 650 655

Gln Glu Pro Phe Gln Leu Ala Thr Glu Lys Arg Ala Lys Glu Arg Gln  
660 665 670

Glu Leu Glu Lys Arg Met Ala Glu Val Glu Ala Gln Lys Ala Gln Gln  
675 680 685

Leu Glu Glu Ala Arg Leu Gln Glu Glu Glu Gln Lys Lys Glu Glu Leu  
690 695 700

Ala Arg Leu Arg Arg Glu Leu Val His Lys Ala Asn Pro Ile Arg Lys  
705 710 715 720

Tyr Gln Gly Leu Glu Ile Lys Ser Ser Asp Gln Pro Leu Thr Val Pro  
 725 730 735

Val Ser Pro Lys Phe Ser Thr Arg Phe His Cys  
 740 745

<210> 35

<211> 1246

<212> DNA

<213> NM\_001769 CD9

<220>

<221> CDS

<222> (112)..(795)

<223>

<400> 35  
 gaccagccta cagccgcctg catctgtatc cagcgccagg tcccgccagt cccagctgag 60  
 cgcgccccc agtccgcac ccgttcggcc caggctaagt tagccctcac c atg ccg 117  
 Met Pro  
 1  
 gtc aaa gga ggc acc aag tgc atc aaa tac ctg ctg ttc gga ttt aac 165  
 Val Lys Gly Gly Thr Lys Cys Ile Lys Tyr Leu Leu Phe Gly Phe Asn  
 5 10 15  
 ttc atc ttc tgg ctt gcc ggg att gct gtc ctt gcc att gga cta tgg 213  
 Phe Ile Phe Trp Leu Ala Gly Ile Ala Val Leu Ala Ile Gly Leu Trp  
 20 25 30  
 ctc cga ttc gac tct cag acc aag agc atc ttc gag caa gaa act aat 261  
 Leu Arg Phe Asp Ser Gln Thr Lys Ser Ile Phe Glu Gln Glu Thr Asn  
 35 40 45 50  
 aat aat aat tcc agc ttc tac aca gga gtc tat att ctg atc gga gcc 309  
 Asn Asn Asn Ser Ser Phe Tyr Thr Gly Val Tyr Ile Leu Ile Gly Ala  
 55 60 65  
 ggc gcc ctc atg atg ctg gtg ggc ttc ctg ggc tgc tgc ggg gct gtg 357  
 Gly Ala Leu Met Met Leu Val Gly Phe Leu Gly Cys Cys Gly Ala Val  
 70 75 80  
 cag gag tcc cag tgc atg ctg gga ctg ttc ttc ggc ttc ctc ttg gtg 405  
 Gln Glu Ser Gln Cys Met Leu Gly Leu Phe Phe Gly Phe Leu Leu Val  
 85 90 95  
 ata ttc gcc att gaa ata gct gcg gcc atc tgg gga tat tcc cac aag 453  
 Ile Phe Ala Ile Glu Ile Ala Ala Ala Ile Trp Gly Tyr Ser His Lys  
 100 105 110  
 gat gag gtg att aag gaa gtc cag gag ttt tac aag gac acc tac aac 501  
 Asp Glu Val Ile Lys Glu Val Gln Glu Phe Tyr Lys Asp Thr Tyr Asn  
 115 120 125 130  
 aag ctg aaa acc aag gat gag ccc cag cgg gaa acg ctg aaa gcc atc 549



Lys Leu Lys Thr Lys Asp Glu Pro Gln Arg Glu Thr Leu Lys Ala Ile  
 135 140 145  
 cac tat gcg ttg aac tgc tgt ggt ttg gct ggg ggc gtg gaa cag ttt 597  
 His Tyr Ala Leu Asn Cys Cys Gly Leu Ala Gly Gly Val Glu Gln Phe  
 150 155 160  
 atc tca gac atc tgc ccc aag aag gac gta ctc gaa acc ttc acc gtg 645  
 Ile Ser Asp Ile Cys Pro Lys Lys Asp Val Leu Glu Thr Phe Thr Val  
 165 170 175  
 aag tcc tgt cct gat gcc atc aaa gag gtc ttc gac aat aaa ttc cac 693  
 Lys Ser Cys Pro Asp Ala Ile Lys Glu Val Phe Asp Asn Lys Phe His  
 180 185 190  
 atc atc ggc gca gtg ggc atc ggc att gcc gtg gtc atg ata ttt ggc 741  
 Ile Ile Gly Ala Val Gly Ile Gly Ile Ala Val Val Met Ile Phe Gly  
 195 200 205 210  
 atg atc ttc agt atg atc ttg tgc tgt gct atc cgc agg aac cgc gag 789  
 Met Ile Phe Ser Met Ile Leu Cys Cys Ala Ile Arg Arg Asn Arg Glu  
 215 220 225  
 atg gtc tagagtcagc ttacatccct gagcaggaaa gtttaccat gaagattggt 845  
 Met Val  
 gggatttttt gtttgtttgt tttgttttgt ttgttgtttg ttgtttgttt ttttgccact 905  
 aatttttagta ttcatctctgc attgctagat aaaagctgaa gttactttat gtttgtcttt 965  
 taatgcttca ttcaatattg acatttgtag ttgagcgggg gttttggttt gctttggttt 1025  
 atattttttc agttgtttgt ttttgcttgt tatattaagc agaaatcctg caatgaaagg 1085  
 tactatatit gctagactct agacaagata ttgtacataa aagaattttt ttgtctttta 1145  
 atagatacaa atgtctatca actttaatca agttgtaact tatattgaag acaatttgat 1205  
 acataataaa aaattatgac aatgtcaaaa aaaaaaaaaa a 1246  
 <210> 36  
 <211> 228  
 <212> PRT  
 <213> NM\_001769 CD9  
 <400> 36  
 Met Pro Val Lys Gly Gly Thr Lys Cys Ile Lys Tyr Leu Leu Phe Gly  
 1 5 10 15  
 Phe Asn Phe Ile Phe Trp Leu Ala Gly Ile Ala Val Leu Ala Ile Gly  
 20 25 30  
 Leu Trp Leu Arg Phe Asp Ser Gln Thr Lys Ser Ile Phe Glu Gln Glu  
 35 40 45  
 Thr Asn Asn Asn Asn Ser Ser Phe Tyr Thr Gly Val Tyr Ile Leu Ile  
 50 55 60

Gly Ala Gly Ala Leu Met Met Leu Val Gly Phe Leu Gly Cys Cys Gly  
65 70 75 80

Ala Val Gln Glu Ser Gln Cys Met Leu Gly Leu Phe Phe Gly Phe Leu  
85 90 95

Leu Val Ile Phe Ala Ile Glu Ile Ala Ala Ala Ile Trp Gly Tyr Ser  
100 105 110

His Lys Asp Glu Val Ile Lys Glu Val Gln Glu Phe Tyr Lys Asp Thr  
115 120 125

Tyr Asn Lys Leu Lys Thr Lys Asp Glu Pro Gln Arg Glu Thr Leu Lys  
130 135 140

Ala Ile His Tyr Ala Leu Asn Cys Cys Gly Leu Ala Gly Gly Val Glu  
145 150 155 160

Gln Phe Ile Ser Asp Ile Cys Pro Lys Lys Asp Val Leu Glu Thr Phe  
165 170 175

Thr Val Lys Ser Cys Pro Asp Ala Ile Lys Glu Val Phe Asp Asn Lys  
180 185 190

Phe His Ile Ile Gly Ala Val Gly Ile Gly Ile Ala Val Val Met Ile  
195 200 205

Phe Gly Met Ile Phe Ser Met Ile Leu Cys Cys Ala Ile Arg Arg Asn  
210 215 220

Arg Glu Met Val  
225

<210> 37

<211> 7557

<212> DNA

<213> NM\_020859 ShrmL, Shroom-related protein

<220>

<221> CDS

<222> (958) .. (6942)

<223>

<400> 37

ggttattaac tgtaagactg tttgaattgt caggtctttg tggtagactg aattctgcat

60

|  |      |
|--|------|
| ggcagtgctt actgtgcaga aagggttcaa tgcctcactt tgccttaatc catgttcctc  | 120  |
| taacttcctt actctttctc tcaagaggag gcaagtggct gtggcggcgg cagcagtggc  | 180  |
| tgatcatcac tgaaaatacc aaagaaaaga actgagctgc ctcttcata tttttccat    | 240  |
| tgaggattaa ttaccgtgc tttttcattt tctctacatc ctgcaaaagt ttttttctct   | 300  |
| cctaagaaac aaactatgaa ctgattgttg aaaaaaagaa gtaaaaagtt ttagcacagc  | 360  |
| ttctctgtct cttcgggaca agttagaaaa ttctgaagtg agccgaagca tagtaagtgc  | 420  |
| tttctttctt ttttaagctac ttctggggag ggaggaggct attgtaatgg taaatttcac | 480  |
| tccgagagga agaaagggtt gataatcaat caaaaatgag gtatttcctt gagtatttgt  | 540  |
| gattttctta ctatattgta agatgctttt aattttctct gtaaaatagg cagaaatgg   | 600  |
| tttagtgtgt gtatgtgtga aataaaagct cagaaaagca atcttcagag cgccactgaa  | 660  |
| ggaagttttg acgaacggag tagagatgta taccacttgg gggcttcagt gagaaccag   | 720  |
| aattcctgga ggaggattta cattcagaaa tgttgaagtg aaaattcctt ctggttcagc  | 780  |
| atcttggagt tcagcttgga agaacatttt acgtatggaa gaatttgctt ctccaaacct  | 840  |
| ctcttttggc cattgtgtgt cctgaaggat gggacaactt gtgctgtaga agcactgctt  | 900  |
| gcctgagttt gcttcaggca ttttaaattt aacttgaggg atcatgtgtt tggcatg     | 957  |
| atg agg acc act gaa gac ttc cac aag cct agt gcc aca tta aac tct    | 1005 |
| Met Arg Thr Thr Glu Asp Phe His Lys Pro Ser Ala Thr Leu Asn Ser    |      |
| 1 5 10 15  |      |
| aac acg gcc acc aag gga agg tac att tat ctg gag gca ttc ctg gag    | 1053 |
| Asn Thr Ala Thr Lys Gly Arg Tyr Ile Tyr Leu Glu Ala Phe Leu Glu    |      |
| 20 25 30   |      |
| gga gga gct ccc tgg ggt ttt act cta aag ggt ggc ctg gag cac gga    | 1101 |
| Gly Gly Ala Pro Trp Gly Phe Thr Leu Lys Gly Gly Leu Glu His Gly    |      |
| 35 40 45   |      |
| gaa cca tta atc atc tct aag gtc gaa gaa ggg ggc aaa gca gac acc    | 1149 |
| Glu Pro Leu Ile Ile Ser Lys Val Glu Glu Gly Gly Lys Ala Asp Thr    |      |
| 50 55 60   |      |
| ctg agc tcc aaa ctg cag gct ggg gat gag gtt gtg cac atc aat gag    | 1197 |
| Leu Ser Ser Lys Leu Gln Ala Gly Asp Glu Val Val His Ile Asn Glu    |      |
| 65 70 75 80  |      |
| gtg act ctg agc agc tcc aga aag gag gca gtt tcc ctg gtg aaa gga    | 1245 |
| Val Thr Leu Ser Ser Arg Lys Glu Ala Val Ser Leu Val Lys Gly        |      |
| 85 90 95   |      |
| tcc tac aag acc ctc agg ctg gta gtg cgc aga gat gtg tgc aca gac    | 1293 |
| Ser Tyr Lys Thr Leu Arg Leu Val Arg Arg Asp Val Cys Thr Asp        |      |
| 100 105 110  |      |
| cca ggc cat gca gat act ggt gcc tct aac ttc gtc agc cca gaa cac    | 1341 |
| Pro Gly His Ala Asp Thr Gly Ala Ser Asn Phe Val Ser Pro Glu His    |      |
| 115 120 125  |      |
| ctc acc tct ggc ccc cag cac agg aaa gca gcg tgg tca gga ggg gtt    | 1389 |
| Leu Thr Ser Gly Pro Gln His Arg Lys Ala Ala Trp Ser Gly Gly Val    |      |
| 130 135 140  |      |
| aaa ctt cgg ctg aag cac agg tct agt gag cct gca ggc cga cct cac    | 1437 |

The authors are grateful to the referees for their valuable comments and suggestions. The authors also thank the Department of Mathematics, University of Jammu, for providing the facilities for carrying out this work.

|   |      |
|---|------|
| Asn Ser Leu Gly Ser Leu Lys Ser Pro Phe Ile Glu Glu Gln Leu His |      |
| 420 425 430   |      |
| act gtg ctg gag aag agt cca gag aac agc ccc cca gtg aag ccc aag | 2301 |
| Thr Val Leu Glu Lys Ser Pro Glu Asn Ser Pro Pro Val Lys Pro Lys |      |
| 435 440 445   |      |
| cat aac tat acc cag aag gcc caa cct gcc caa cct ctg ctg ccg acc | 2349 |
| His Asn Tyr Thr Gln Lys Ala Gln Pro Gly Gln Pro Leu Leu Pro Thr |      |
| 450 455 460   |      |
| agc atc tac gcg gta cct tcc ctg gag cca cac ttt gcc cag gtg cct | 2397 |
| Ser Ile Tyr Ala Val Pro Ser Leu Glu Pro His Phe Ala Gln Val Pro |      |
| 465 470 475 480   |      |
| cag cct tct gtg agt agc aac ggt atg ctc tac cct gca ctg gcc aag | 2445 |
| Gln Pro Ser Val Ser Ser Asn Gly Met Leu Tyr Pro Ala Leu Ala Lys |      |
| 485 490 495   |      |
| gag agt gga tac ata gcc cct cag gga gca tgc aac aag atg gct acc | 2493 |
| Glu Ser Gly Tyr Ile Ala Pro Gln Gly Ala Cys Asn Lys Met Ala Thr |      |
| 500 505 510   |      |
| att gat gag aat ggg aac cag aat gga tct gcc agg cct ggg ttt gcc | 2541 |
| Ile Asp Glu Asn Gly Asn Gln Asn Gly Ser Gly Arg Pro Gly Phe Ala |      |
| 515 520 525   |      |
| ttc tgc cag ccc tta gaa cat gac ttg ctg tcc cca gtg gag aag aaa | 2589 |
| Phe Cys Gln Pro Leu Glu His Asp Leu Leu Ser Pro Val Glu Lys Lys |      |
| 530 535 540   |      |
| cca gaa gct aca gcc aag tat gtc ccc tcc aaa gtc cat ttc tgt tca | 2637 |
| Pro Glu Ala Thr Ala Lys Tyr Val Pro Ser Lys Val His Phe Cys Ser |      |
| 545 550 555 560   |      |
| gtg cct gaa aat gag gag gat gcc tcc ctg aag aga cat ctc aca cct | 2685 |
| Val Pro Glu Asn Glu Glu Asp Ala Ser Leu Lys Arg His Leu Thr Pro |      |
| 565 570 575   |      |
| ccc caa ggc aac agc cca cat tcc aat gag aga aag agc acc cac agt | 2733 |
| Pro Gln Gly Asn Ser Pro His Ser Asn Glu Arg Lys Ser Thr His Ser |      |
| 580 585 590   |      |
| aac aaa cca tct tct cat ccc cac agc ctc aaa tgc cct cag gct cag | 2781 |
| Asn Lys Pro Ser His Pro His Ser Leu Lys Cys Pro Gln Ala Gln     |      |
| 595 600 605   |      |
| gcc tgg caa gcg ggt gaa gac aag aga tct tcc agg ctc tca gag ccc | 2829 |
| Ala Trp Gln Ala Gly Glu Asp Lys Arg Ser Ser Arg Leu Ser Glu Pro |      |
| 610 615 620   |      |
| tgg gag ggc gat ttc cag gaa gac cac aat gcc aac ctc tgg agg agg | 2877 |
| Trp Glu Gly Asp Phe Gln Glu Asp His Asn Ala Asn Leu Trp Arg Arg |      |
| 625 630 635 640   |      |
| ctg gag aga gaa ggc cta ggc cag agc ctg tca ggc aac ttt ggc aag | 2925 |
| Leu Glu Arg Glu Gly Leu Gly Gln Ser Leu Ser Gly Asn Phe Gly Lys |      |
| 645 650 655   |      |
| acc aag tca gcc ttc tca tct ctc cag aac att cct gag agt ctg aga | 2973 |
| Thr Lys Ser Ala Phe Ser Ser Leu Gln Asn Ile Pro Glu Ser Leu Arg |      |
| 660 665 670   |      |
| aga cac agc agc ctg gag cta ggc cgg gga acc cag gag ggt tac ccc | 3021 |
| Arg His Ser Ser Leu Glu Leu Gly Arg Gly Thr Gln Glu Gly Tyr Pro |      |
| 675 680 685   |      |
| ggg ggc agg ccc acc tgt gca gtc aac acc aag gca gaa gac cct ggg | 3069 |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| Gly | Gly | Arg | Pro | Thr | Cys | Ala | Val | Asn | Thr | Lys | Ala | Glu | Asp | Pro | Gly |      |
| 690 |     |     |     |     |     | 695 |     |     |     |     | 700 |     |     |     |     |      |
| agg | aaa | gcc | gct | cct | gac | ctc | ggg | agc | cat | ctg | gac | cgg | cag | gtt | tcc | 3117 |
| Arg | Lys | Ala | Ala | Pro | Asp | Leu | Gly | Ser | His | Leu | Asp | Arg | Gln | Val | Ser |      |
| 705 |     |     |     |     |     | 710 |     |     |     | 715 |     |     |     |     | 720 |      |
| tac | ccg | cgg | ccc | gag | ggg | agg | acc | ggt | gcc | tcg | gct | tct | ttc | aac | agc | 3165 |
| Tyr | Pro | Arg | Pro | Glu | Gly | Arg | Thr | Gly | Ala | Ser | Ala | Ser | Phe | Asn | Ser |      |
|     |     |     |     | 725 |     |     |     |     | 730 |     |     |     |     | 735 |     |      |
| aca | gac | cca | agt | ccc | gaa | gag | ccg | cct | gcc | ccc | tcg | cac | ccg | cac | aca | 3213 |
| Thr | Asp | Pro | Ser | Pro | Glu | Glu | Pro | Pro | Ala | Pro | Ser | His | Pro | His | Thr |      |
|     |     |     | 740 |     |     |     |     | 745 |     |     |     |     | 750 |     |     |      |
| tcc | agt | ctg | ggc | cgg | agg | ggg | ccc | ggc | cca | ggc | agc | gcc | tcg | gct | ctt | 3261 |
| Ser | Ser | Leu | Gly | Arg | Arg | Gly | Pro | Gly | Pro | Gly | Ser | Ala | Ser | Ala | Leu |      |
|     |     | 755 |     |     |     |     | 760 |     |     |     |     | 765 |     |     |     |      |
| cag | ggc | ttt | cag | tac | ggg | aag | ccc | cac | tgc | tcg | gtg | ctg | gag | aag | gtc | 3309 |
| Gln | Gly | Phe | Gln | Tyr | Gly | Lys | Pro | His | Cys | Ser | Val | Leu | Glu | Lys | Val |      |
|     |     | 770 |     |     |     | 775 |     |     |     |     | 780 |     |     |     |     |      |
| tcc | aaa | ttc | gag | cag | cga | gag | caa | ggg | agc | cag | aga | ccg | agt | gtg | ggc | 3357 |
| Ser | Lys | Phe | Glu | Gln | Arg | Glu | Gln | Gly | Ser | Gln | Arg | Pro | Ser | Val | Gly |      |
|     |     |     |     |     |     | 790 |     |     |     | 795 |     |     |     |     | 800 |      |
| ggc | tct | ggt | ttt | ggc | cat | aac | tat | agg | ccc | cac | agg | acc | gtc | tca | act | 3405 |
| Gly | Ser | Gly | Phe | Gly | His | Asn | Tyr | Arg | Pro | His | Arg | Thr | Val | Ser | Thr |      |
|     |     |     |     | 805 |     |     |     |     | 810 |     |     |     |     | 815 |     |      |
| tcc | agt | act | tct | ggg | aat | gac | ttc | gag | gag | aca | aaa | gca | cac | att | cgt | 3453 |
| Ser | Ser | Thr | Ser | Gly | Asn | Asp | Phe | Glu | Glu | Thr | Lys | Ala | His | Ile | Arg |      |
|     |     |     | 820 |     |     |     |     | 825 |     |     |     |     | 830 |     |     |      |
| ttc | tct | gag | tca | gct | gaa | ccc | cta | ggc | aac | ggg | gag | cag | cac | ttc | aaa | 3501 |
| Phe | Ser | Glu | Ser | Ala | Glu | Pro | Leu | Gly | Asn | Gly | Glu | Gln | His | Phe | Lys |      |
|     |     | 835 |     |     |     |     | 840 |     |     |     |     | 845 |     |     |     |      |
| aac | ggg | gag | ctg | aag | ttg | gaa | gag | gct | tcc | cgg | cag | ccc | tgc | ggt | cag | 3549 |
| Asn | Gly | Glu | Leu | Lys | Leu | Glu | Glu | Ala | Ser | Arg | Gln | Pro | Cys | Gly | Gln |      |
|     |     | 850 |     |     |     | 855 |     |     |     |     | 860 |     |     |     |     |      |
| cag | ctg | agc | gga | gga | gcg | tcg | gac | agc | ggc | cgt | ggc | ccc | cag | agg | ccg | 3597 |
| Gln | Leu | Ser | Gly | Gly | Ala | Ser | Asp | Ser | Gly | Arg | Gly | Pro | Gln | Arg | Pro |      |
|     |     |     |     |     | 870 |     |     |     | 875 |     |     |     |     | 880 |     |      |
| gac | gct | cgg | ctc | ctc | cgt | agc | cag | agc | acc | ttc | cag | ctc | tcc | agc | gag | 3645 |
| Asp | Ala | Arg | Leu | Leu | Arg | Ser | Gln | Ser | Thr | Phe | Gln | Leu | Ser | Ser | Glu |      |
|     |     |     |     | 885 |     |     |     |     | 890 |     |     |     |     | 895 |     |      |
| cca | gag | agg | gag | ccc | gag | tgg | cgg | gac | agg | ccc | ggc | tcg | ccc | gaa | tcg | 3693 |
| Pro | Glu | Arg | Glu | Pro | Glu | Trp | Arg | Asp | Arg | Pro | Gly | Ser | Pro | Glu | Ser |      |
|     |     |     | 900 |     |     |     |     | 905 |     |     |     |     | 910 |     |     |      |
| ccc | ctg | ctg | gat | gcc | ccc | ttc | agc | cgc | gcc | tac | cgg | aac | agc | atc | aag | 3741 |
| Pro | Leu | Leu | Asp | Ala | Pro | Phe | Ser | Arg | Ala | Tyr | Arg | Asn | Ser | Ile | Lys |      |
|     |     | 915 |     |     |     |     | 920 |     |     |     |     | 925 |     |     |     |      |
| gac | gca | cag | tcc | cgt | gtc | ttg | ggg | gcc | acc | tcc | ttt | cga | cgt | cga | gac | 3789 |
| Asp | Ala | Gln | Ser | Arg | Val | Leu | Gly | Ala | Thr | Ser | Phe | Arg | Arg | Arg | Asp |      |
|     |     | 930 |     |     |     | 935 |     |     |     |     | 940 |     |     |     |     |      |
| ctg | gag | ctg | ggg | gcg | ccc | gtg | gcg | tcg | agg | tcc | tgg | cgg | cca | cgg | cct | 3837 |
| Leu | Glu | Leu | Gly | Ala | Pro | Val | Ala | Ser | Arg | Ser | Trp | Arg | Pro | Arg | Pro |      |
|     |     | 945 |     |     | 950 |     |     |     | 955 |     |     |     |     |     | 960 |      |
| tcc | tcg | gcc | cac | gtg | ggg | ctg | cgg | agc | ccc | gag | gcg | tcg | gcc | tcc | gcc | 3885 |

|     |      |     |     |     |     |      |      |     |     |     |      |     |      |     |     |      |
|-----|------|-----|-----|-----|-----|------|------|-----|-----|-----|------|-----|------|-----|-----|------|
| Ser | Ser  | Ala | His | Val | Gly | Leu  | Arg  | Ser | Pro | Glu | Ala  | Ser | Ala  | Ser | Ala |      |
|     |      |     |     | 965 |     |      |      |     | 970 |     |      |     |      |     | 975 |      |
| tcc | ccg  | cac | acg | ccc | cgg | gag  | cgg  | cac | agc | gtg | acc  | cct | gct  | gag | ggc | 3933 |
| Ser | Pro  | His | Thr | Pro | Arg | Glu  | Arg  | His | Ser | Val | Thr  | Pro | Ala  | Glu | Gly |      |
|     |      |     | 980 |     |     |      |      | 985 |     |     |      |     |      | 990 |     |      |
| gac | ctg  | gcc | agg | ccc | gtg | ccc  | cct  | gcc | gcc | cgg | aga  | ggt | gct  | cgc | cgg | 3981 |
| Asp | Leu  | Ala | Arg | Pro | Val | Pro  | Pro  | Ala | Ala | Arg | Arg  | Gly | Ala  | Arg | Arg |      |
|     |      |     | 995 |     |     |      | 1000 |     |     |     |      |     | 1005 |     |     |      |
| cgc | ctg  | act | ccc | gag | cag | aag  | aag  | cgc | tcc | tac | tcg  | gag | ccc  | gag |     | 4026 |
| Arg | Leu  | Thr | Pro | Glu | Gln | Lys  | Lys  | Arg | Ser | Tyr | Ser  | Glu | Pro  | Glu |     |      |
|     | 1010 |     |     |     |     | 1015 |      |     |     |     | 1020 |     |      |     |     |      |
| aag | atg  | aac | gag | gtg | ggg | atc  | gtg  | gag | gag | gcc | gaa  | ccg | gca  | ccc |     | 4071 |
| Lys | Met  | Asn | Glu | Val | Gly | Ile  | Val  | Glu | Glu | Ala | Glu  | Pro | Ala  | Pro |     |      |
|     | 1025 |     |     |     |     | 1030 |      |     |     |     | 1035 |     |      |     |     |      |
| ctg | ggc  | ccg | cag | aga | aat | ggg  | atg  | cgt | ttc | ccg | gag  | agc | agc  | gtg |     | 4116 |
| Leu | Gly  | Pro | Gln | Arg | Asn | Gly  | Met  | Arg | Phe | Pro | Glu  | Ser | Ser  | Val |     |      |
|     | 1040 |     |     |     |     | 1045 |      |     |     |     | 1050 |     |      |     |     |      |
| gcc | gac  | cgg | cgc | cgt | ctc | ttc  | gag  | cgc | gat | ggc | aag  | gcc | tgc  | tcc |     | 4161 |
| Ala | Asp  | Arg | Arg | Arg | Leu | Phe  | Glu  | Arg | Asp | Gly | Lys  | Ala | Cys  | Ser |     |      |
|     | 1055 |     |     |     |     | 1060 |      |     |     |     | 1065 |     |      |     |     |      |
| acg | ctc  | agc | ctg | tcg | ggg | ccc  | gag  | ctg | aag | cag | ttc  | cag | cag  | agc |     | 4206 |
| Thr | Leu  | Ser | Leu | Ser | Gly | Pro  | Glu  | Leu | Lys | Gln | Phe  | Gln | Gln  | Ser |     |      |
|     | 1070 |     |     |     |     | 1075 |      |     |     |     | 1080 |     |      |     |     |      |
| gcc | ctg  | gcg | gac | tac | atc | cag  | cgc  | aag | acc | ggc | aag  | cgg | cct  | acc |     | 4251 |
| Ala | Leu  | Ala | Asp | Tyr | Ile | Gln  | Arg  | Lys | Thr | Gly | Lys  | Arg | Pro  | Thr |     |      |
|     | 1085 |     |     |     |     | 1090 |      |     |     |     | 1095 |     |      |     |     |      |
| tcc | gcc  | gcc | ggc | tgc | agc | ctc  | cag  | gag | ccc | ggg | cca  | ctg | cgt  | gag |     | 4296 |
| Ser | Ala  | Ala | Gly | Cys | Ser | Leu  | Gln  | Glu | Pro | Gly | Pro  | Leu | Arg  | Glu |     |      |
|     | 1100 |     |     |     |     | 1105 |      |     |     |     | 1110 |     |      |     |     |      |
| cgc | gcc  | cag | agt | gcc | tac | ctc  | cag  | ccc | ggc | ccc | gcg  | gcg | ctc  | gaa |     | 4341 |
| Arg | Ala  | Gln | Ser | Ala | Tyr | Leu  | Gln  | Pro | Gly | Pro | Ala  | Ala | Leu  | Glu |     |      |
|     | 1115 |     |     |     |     | 1120 |      |     |     |     | 1125 |     |      |     |     |      |
| ggc | tcc  | ggc | ctc | gcc | tcg | gcc  | tcc  | agc | ttg | agc | tca  | ctg | cgg  | gag |     | 4386 |
| Gly | Ser  | Gly | Leu | Ala | Ser | Ala  | Ser  | Ser | Leu | Ser | Ser  | Leu | Arg  | Glu |     |      |
|     | 1130 |     |     |     |     | 1135 |      |     |     |     | 1140 |     |      |     |     |      |
| ccc | agc  | ctg | cag | ccc | cgc | agg  | gag  | gcc | acg | ctc | ctg  | ccg | gcc  | aca |     | 4431 |
| Pro | Ser  | Leu | Gln | Pro | Arg | Arg  | Glu  | Ala | Thr | Leu | Leu  | Pro | Ala  | Thr |     |      |
|     | 1145 |     |     |     |     | 1150 |      |     |     |     | 1155 |     |      |     |     |      |
| gtt | gca  | gaa | acc | cag | cag | gct  | ccc  | cga | gat | cgc | agc  | agc | tcc  | ttc |     | 4476 |
| Val | Ala  | Glu | Thr | Gln | Gln | Ala  | Pro  | Arg | Asp | Arg | Ser  | Ser | Ser  | Phe |     |      |
|     | 1160 |     |     |     |     | 1165 |      |     |     |     | 1170 |     |      |     |     |      |
| gcc | ggt  | ggc | cgc | cgc | ctc | ggg  | gaa  | cgg | cga | cgc | ggg  | gac | ctg  | ctt |     | 4521 |
| Ala | Gly  | Gly | Arg | Arg | Leu | Gly  | Glu  | Arg | Arg | Arg | Gly  | Asp | Leu  | Leu |     |      |
|     | 1175 |     |     |     |     | 1180 |      |     |     |     | 1185 |     |      |     |     |      |
| agc | gga  | gca | aac | ggt | gga | aca  | agg  | ggc | acc | cag | aga  | ggg | gat  | gag |     | 4566 |
| Ser | Gly  | Ala | Asn | Gly | Gly | Thr  | Arg  | Gly | Thr | Gln | Arg  | Gly | Asp  | Glu |     |      |
|     | 1190 |     |     |     |     | 1195 |      |     |     |     | 1200 |     |      |     |     |      |
| acc | ccc  | agg | gag | cca | tcc | tcc  | tgg  | ggg | gcc | agg | gcc  | ggg | aag  | tcc |     | 4611 |
| Thr | Pro  | Arg | Glu | Pro | Ser | Ser  | Trp  | Gly | Ala | Arg | Ala  | Gly | Lys  | Ser |     |      |
|     | 1205 |     |     |     |     | 1210 |      |     |     |     | 1215 |     |      |     |     |      |
| atg | tcg  | gcc | gag | gac | ctg | ctg  | gaa  | cgc | tcg | gac | gtc  | ctt | gcg  | ggc |     | 4656 |

|      |     |     |     |     |     |      |     |     |     |     |      |     |     |     |      |
|------|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|------|-----|-----|-----|------|
| Met  | Ser | Ala | Glu | Asp | Leu | Leu  | Glu | Arg | Ser | Asp | Val  | Leu | Ala | Gly |      |
| 1220 |     |     |     |     |     | 1225 |     |     |     |     | 1230 |     |     |     |      |
| cct  | gtc | cat | gtg | agg | tcc | agg  | tca | tct | ccc | gcc | acc  | gca | gac | aag | 4701 |
| Pro  | Val | His | Val | Arg | Ser | Arg  | Ser | Ser | Pro | Ala | Thr  | Ala | Asp | Lys |      |
| 1235 |     |     |     |     |     | 1240 |     |     |     |     | 1245 |     |     |     |      |
| cgc  | cag | gat | gtg | ctt | ttg | ggg  | caa | gac | agt | ggc | ttt  | ggg | ctt | gtg | 4746 |
| Arg  | Gln | Asp | Val | Leu | Leu | Gly  | Gln | Asp | Ser | Gly | Phe  | Gly | Leu | Val |      |
| 1250 |     |     |     |     |     | 1255 |     |     |     |     | 1260 |     |     |     |      |
| aag  | gat | cca | tgt | tat | ttg | gct  | ggg | cct | gga | tct | agg  | tca | ctc | agt | 4791 |
| Lys  | Asp | Pro | Cys | Tyr | Leu | Ala  | Gly | Pro | Gly | Ser | Arg  | Ser | Leu | Ser |      |
| 1265 |     |     |     |     |     | 1270 |     |     |     |     | 1275 |     |     |     |      |
| tgt  | tca | gaa | aga | ggc | caa | gaa  | gag | atg | ctg | ctg | ctc  | ttc | cac | cat | 4836 |
| Cys  | Ser | Glu | Arg | Gly | Gln | Glu  | Glu | Met | Leu | Leu | Leu  | Phe | His | His |      |
| 1280 |     |     |     |     |     | 1285 |     |     |     |     | 1290 |     |     |     |      |
| ctc  | acc | cct | cgt | tgg | ggg | ggg  | tca | ggc | tgc | aaa | gcc  | att | ggg | gat | 4881 |
| Leu  | Thr | Pro | Arg | Trp | Gly | Gly  | Ser | Gly | Cys | Lys | Ala  | Ile | Gly | Asp |      |
| 1295 |     |     |     |     |     | 1300 |     |     |     |     | 1305 |     |     |     |      |
| tcc  | tcc | gtt | cct | agt | gaa | tgt  | cct | gga | acc | ctg | gac  | cat | cag | agg | 4926 |
| Ser  | Ser | Val | Pro | Ser | Glu | Cys  | Pro | Gly | Thr | Leu | Asp  | His | Gln | Arg |      |
| 1310 |     |     |     |     |     | 1315 |     |     |     |     | 1320 |     |     |     |      |
| caa  | gcc | agt | agg | aca | ccc | tgc  | ccc | agg | cca | cca | ctg  | gca | gga | acg | 4971 |
| Gln  | Ala | Ser | Arg | Thr | Pro | Cys  | Pro | Arg | Pro | Pro | Leu  | Ala | Gly | Thr |      |
| 1325 |     |     |     |     |     | 1330 |     |     |     |     | 1335 |     |     |     |      |
| caa  | ggg | ctg | gtc | aca | gac | acc  | agg | gct | gca | ccc | ctg  | acc | cca | att | 5016 |
| Gln  | Gly | Leu | Val | Thr | Asp | Thr  | Arg | Ala | Ala | Pro | Leu  | Thr | Pro | Ile |      |
| 1340 |     |     |     |     |     | 1345 |     |     |     |     | 1350 |     |     |     |      |
| ggc  | acc | cct | ctg | cct | tca | gcc  | att | ccc | tct | ggc | tac  | tgc | tca | cag | 5061 |
| Gly  | Thr | Pro | Leu | Pro | Ser | Ala  | Ile | Pro | Ser | Gly | Tyr  | Cys | Ser | Gln |      |
| 1355 |     |     |     |     |     | 1360 |     |     |     |     | 1365 |     |     |     |      |
| gac  | ggg | cag | aca | ggg | cga | cag  | cct | ctc | ccg | ccc | tac  | acc | cct | gcc | 5106 |
| Asp  | Gly | Gln | Thr | Gly | Arg | Gln  | Pro | Leu | Pro | Pro | Tyr  | Thr | Pro | Ala |      |
| 1370 |     |     |     |     |     | 1375 |     |     |     |     | 1380 |     |     |     |      |
| atg  | atg | cac | aga | agc | aat | ggg  | cac | acc | ctg | acc | cag  | cct | ccc | ggg | 5151 |
| Met  | Met | His | Arg | Ser | Asn | Gly  | His | Thr | Leu | Thr | Gln  | Pro | Pro | Gly |      |
| 1385 |     |     |     |     |     | 1390 |     |     |     |     | 1395 |     |     |     |      |
| cca  | aga | ggc | tgt | gag | ggc | gat  | ggc | cca | gag | cat | ggg  | gta | gaa | gag | 5196 |
| Pro  | Arg | Gly | Cys | Glu | Gly | Asp  | Gly | Pro | Glu | His | Gly  | Val | Glu | Glu |      |
| 1400 |     |     |     |     |     | 1405 |     |     |     |     | 1410 |     |     |     |      |
| gga  | acg | agg | aag | agg | gtc | tgc  | ctg | cct | cag | tgg | cca  | cct | cct | tct | 5241 |
| Gly  | Thr | Arg | Lys | Arg | Val | Ser  | Leu | Pro | Gln | Trp | Pro  | Pro | Pro | Ser |      |
| 1415 |     |     |     |     |     | 1420 |     |     |     |     | 1425 |     |     |     |      |
| cga  | gca | aag | tgg | gcc | cac | gca  | gcc | aga | gag | gac | agc  | ctt | cct | gag | 5286 |
| Arg  | Ala | Lys | Trp | Ala | His | Ala  | Ala | Arg | Glu | Asp | Ser  | Leu | Pro | Glu |      |
| 1430 |     |     |     |     |     | 1435 |     |     |     |     | 1440 |     |     |     |      |
| gaa  | tcc | tca | gcc | cct | gat | ttt  | gca | aac | ctg | aag | cac  | tat | caa | aaa | 5331 |
| Glu  | Ser | Ser | Ala | Pro | Asp | Phe  | Ala | Asn | Leu | Lys | His  | Tyr | Gln | Lys |      |
| 1445 |     |     |     |     |     | 1450 |     |     |     |     | 1455 |     |     |     |      |
| cag  | cag | agt | ctt | cca | agt | tta  | tgc | agc | act | tct | gac  | cca | gac | aca | 5376 |
| Gln  | Gln | Ser | Leu | Pro | Ser | Leu  | Cys | Ser | Thr | Ser | Asp  | Pro | Asp | Thr |      |
| 1460 |     |     |     |     |     | 1465 |     |     |     |     | 1470 |     |     |     |      |
| cct  | ctt | ggg | gcc | ccg | agc | act  | cca | ggg | agg | atc | tcc  | ctc | cga | ata | 5421 |



|                             |                     |             |      |
|-----------------------------|---------------------|-------------|------|
| Pro Leu Gly Ala Pro Ser Thr | Pro Gly Arg Ile Ser | Leu Arg Ile |      |
| 1475                        | 1480                | 1485        |      |
| tct gag tct gtc ctg cgg gac | tcc ccg cca cct cat | gag gat tat | 5466 |
| Ser Glu Ser Val Leu Arg Asp | Ser Pro Pro Pro His | Glu Asp Tyr |      |
| 1490                        | 1495                | 1500        |      |
| gaa gac gaa gtg ttt gtg agg | gat ccg cac ccc aag | gcc acg tcc | 5511 |
| Glu Asp Glu Val Phe Val Arg | Asp Pro His Pro Lys | Ala Thr Ser |      |
| 1505                        | 1510                | 1515        |      |
| agc ccc aca ttt gaa cct ctt | ccc cca ccc cca cct | cct cca ccg | 5556 |
| Ser Pro Thr Phe Glu Pro Leu | Pro Pro Pro Pro Pro | Pro Pro Pro |      |
| 1520                        | 1525                | 1530        |      |
| agt cag gaa acc ccg gtg tat | agc atg gat gac ttc | cct cca cct | 5601 |
| Ser Gln Glu Thr Pro Val Tyr | Ser Met Asp Asp Phe | Pro Pro Pro |      |
| 1535                        | 1540                | 1545        |      |
| cct ccc cac act gta tgt gag | gcg cag ctg gac agt | gag gat ccc | 5646 |
| Pro Pro His Thr Val Cys Glu | Ala Gln Leu Asp Ser | Glu Asp Pro |      |
| 1550                        | 1555                | 1560        |      |
| gag ggg cca cgc ccc agc ttc | aac aaa ctt tct aaa | gtg aca att | 5691 |
| Glu Gly Pro Arg Pro Ser Phe | Asn Lys Leu Ser Lys | Val Thr Ile |      |
| 1565                        | 1570                | 1575        |      |
| gca agg gaa agg cac atg cct | ggt gca gcc cat gtg | gta ggt agt | 5736 |
| Ala Arg Glu Arg His Met Pro | Gly Ala Ala His Val | Val Gly Ser |      |
| 1580                        | 1585                | 1590        |      |
| cag aca ctg gct tcc aga ctc | caa act tct atc aag | ggt tca gag | 5781 |
| Gln Thr Leu Ala Ser Arg Leu | Gln Thr Ser Ile Lys | Gly Ser Glu |      |
| 1595                        | 1600                | 1605        |      |
| gct gag tcc aca cca ccc tcc | ttc atg agc gtt cac | gcc caa ctt | 5826 |
| Ala Glu Ser Thr Pro Pro Ser | Phe Met Ser Val His | Ala Gln Leu |      |
| 1610                        | 1615                | 1620        |      |
| gct ggg tct ctt ggt ggg cag | cca gca ccc atc cag | act caa agc | 5871 |
| Ala Gly Ser Leu Gly Gly Gln | Pro Ala Pro Ile Gln | Thr Gln Ser |      |
| 1625                        | 1630                | 1635        |      |
| ctc agc cat gat cca gtc agt | gga act cag ggt tta | gaa aag aaa | 5916 |
| Leu Ser His Asp Pro Val Ser | Gly Thr Gln Gly Leu | Glu Lys Lys |      |
| 1640                        | 1645                | 1650        |      |
| gtc agt cct gat cct cag aag | agt tca gaa gac atc | aga aca gag | 5961 |
| Val Ser Pro Asp Pro Gln Lys | Ser Ser Glu Asp Ile | Arg Thr Glu |      |
| 1655                        | 1660                | 1665        |      |
| gct ttg gcc aag gaa att gtc | cac caa gac aaa tct | cta gca gac | 6006 |
| Ala Leu Ala Lys Glu Ile Val | His Gln Asp Lys Ser | Leu Ala Asp |      |
| 1670                        | 1675                | 1680        |      |
| att ttg gat cca gac tcc agg | ctg aag aca aca atg | gac ctg atg | 6051 |
| Ile Leu Asp Pro Asp Ser Arg | Leu Lys Thr Thr Met | Asp Leu Met |      |
| 1685                        | 1690                | 1695        |      |
| gaa ggt ttg ttt ccc cga gat | gtg aac ttg ctg aag | gaa aac agt | 6096 |
| Glu Gly Leu Phe Pro Arg Asp | Val Asn Leu Leu Lys | Glu Asn Ser |      |
| 1700                        | 1705                | 1710        |      |
| gta aag agg aag gcc ata cag | aga act gtc agc tct | tca gga tgt | 6141 |
| Val Lys Arg Lys Ala Ile Gln | Arg Thr Val Ser Ser | Ser Gly Cys |      |
| 1715                        | 1720                | 1725        |      |
| gaa ggc aag agg aat gaa gac | aag gaa gca gtg agc | atg ttg gtt | 6186 |

|         |                 |         |                 |         |             |      |             |  |
|---------|-----------------|---------|-----------------|---------|-------------|------|-------------|--|
| Glu Gly | Lys Arg         | Asn Glu | Asp             | Lys Glu | Ala Val     | Ser  | Met Leu Val |  |
| 1730    |                 |         | 1735            |         |             | 1740 |             |  |
| aac tgc | cct gcc tac tac | agt     | gtg tct gct ccc | aag     | gct gag cta | 6231 |             |  |
| Asn Cys | Pro Ala Tyr Tyr | Ser     | Val Ser Ala Pro | Lys     | Ala Glu Leu |      |             |  |
| 1745    |                 | 1750    |                 | 1755    |             |      |             |  |
| ctg aac | aaa atc aaa gag | atg     | cca gca gaa gtg | aat     | gag gaa gag | 6276 |             |  |
| Leu Asn | Lys Ile Lys Glu | Met     | Pro Ala Glu Val | Asn     | Glu Glu Glu |      |             |  |
| 1760    |                 | 1765    |                 | 1770    |             |      |             |  |
| gaa cag | gca gat gtc aat | gaa     | aag aag gct gag | ctc     | att gga agt | 6321 |             |  |
| Glu Gln | Ala Asp Val Asn | Glu     | Lys Lys Ala Glu | Leu     | Ile Gly Ser |      |             |  |
| 1775    |                 | 1780    |                 | 1785    |             |      |             |  |
| ctc acc | cac aag ctg gag | acc     | ctc cag gag gcg | aag     | ggg agc ctg | 6366 |             |  |
| Leu Thr | His Lys Leu Glu | Thr     | Leu Gln Glu Ala | Lys     | Gly Ser Leu |      |             |  |
| 1790    |                 | 1795    |                 | 1800    |             |      |             |  |
| ctc acg | gac atc aag ctc | aac     | aac gcc ctg gga | gaa     | gag gtg gag | 6411 |             |  |
| Leu Thr | Asp Ile Lys Leu | Asn     | Asn Ala Leu Gly | Glu     | Glu Val Glu |      |             |  |
| 1805    |                 | 1810    |                 | 1815    |             |      |             |  |
| gct ctg | atc agc gag ctc | tgc     | aag ccc aat gag | ttt     | gac aag tat | 6456 |             |  |
| Ala Leu | Ile Ser Glu Leu | Cys     | Lys Pro Asn Glu | Phe     | Asp Lys Tyr |      |             |  |
| 1820    |                 | 1825    |                 | 1830    |             |      |             |  |
| agg atg | ttc ata ggg gat | ttg     | gac aag gtg gtc | aac     | ctg ctg ctc | 6501 |             |  |
| Arg Met | Phe Ile Gly Asp | Leu     | Asp Lys Val Val | Asn     | Leu Leu Leu |      |             |  |
| 1835    |                 | 1840    |                 | 1845    |             |      |             |  |
| tcc ctc | tgc ggg cgt cta | gcc     | cgt gtt gag aat | gtc     | ctt agc ggc | 6546 |             |  |
| Ser Leu | Ser Gly Arg Leu | Ala     | Arg Val Glu Asn | Val     | Leu Ser Gly |      |             |  |
| 1850    |                 | 1855    |                 | 1860    |             |      |             |  |
| ctt ggt | gaa gat gcc agt | aat     | gaa gaa agg agc | tct     | ctt tac gag | 6591 |             |  |
| Leu Gly | Glu Asp Ala Ser | Asn     | Glu Glu Arg Ser | Ser     | Leu Tyr Glu |      |             |  |
| 1865    |                 | 1870    |                 | 1875    |             |      |             |  |
| aaa agg | aag atc ctg gct | ggt     | cag cat gag gat | gcc     | cgg gag ctg | 6636 |             |  |
| Lys Arg | Lys Ile Leu Ala | Gly     | Gln His Glu Asp | Ala     | Arg Glu Leu |      |             |  |
| 1880    |                 | 1885    |                 | 1890    |             |      |             |  |
| aag gag | aac ctg gat cgc | agg     | gag cga gta gtg | ctg     | ggc atc ttg | 6681 |             |  |
| Lys Glu | Asn Leu Asp Arg | Arg     | Glu Arg Val Val | Leu     | Gly Ile Leu |      |             |  |
| 1895    |                 | 1900    |                 | 1905    |             |      |             |  |
| gcc aat | tac ctt tca gag | gag     | cag ctc cag gac | tac     | cag cac ttc | 6726 |             |  |
| Ala Asn | Tyr Leu Ser Glu | Glu     | Gln Leu Gln Asp | Tyr     | Gln His Phe |      |             |  |
| 1910    |                 | 1915    |                 | 1920    |             |      |             |  |
| gtg aaa | atg aag tcc acg | ctc     | ctc att gag caa | cgg     | aag ctg gat | 6771 |             |  |
| Val Lys | Met Lys Ser Thr | Leu     | Leu Ile Glu Gln | Arg     | Lys Leu Asp |      |             |  |
| 1925    |                 | 1930    |                 | 1935    |             |      |             |  |
| gac aag | atc aag ctg ggc | cag     | gag cag gtc aag | tgt     | ctg ctg gag | 6816 |             |  |
| Asp Lys | Ile Lys Leu Gly | Gln     | Glu Gln Val Lys | Cys     | Leu Leu Glu |      |             |  |
| 1940    |                 | 1945    |                 | 1950    |             |      |             |  |
| agc ctg | ccc tca gat ttc | att     | ccc aag gct ggg | gcc     | ctg gct ctg | 6861 |             |  |
| Ser Leu | Pro Ser Asp Phe | Ile     | Pro Lys Ala Gly | Ala     | Leu Ala Leu |      |             |  |
| 1955    |                 | 1960    |                 | 1965    |             |      |             |  |
| ccc cca | aac ctc acg agt | gag     | ccc att cct gct | ggg     | ggc tgt act | 6906 |             |  |
| Pro Pro | Asn Leu Thr Ser | Glu     | Pro Ile Pro Ala | Gly     | Gly Cys Thr |      |             |  |
| 1970    |                 | 1975    |                 | 1980    |             |      |             |  |
| ttc agt | ggt att ttc cca | aca     | tta acc tct cca | ctt     | taacctcttc  | 6952 |             |  |

Phe Ser Gly Ile Phe Pro Thr Leu Thr Ser Pro Leu  
 1985 1990 1995

taaaatacc aacaaaaaga tcactgtttc tctcaacact atttaatctg aaaaatgttt 7012  
 cagtacaaac cactgtttga actatctggg ttattggtgt ttgttctga tgaaggaaa 7072  
 aaaattctct ccaggaggaa gcctttttcc ttcttgccct tcttgattga tcttctgaga 7132  
 gctcgaatgc tgctggacac gtaccccttt ctattattac ttgtagtag aaagaaagt 7192  
 aatgaaactg agaactgatt ggagggtgtt tgatcattta gttttaaca ggctgaggca 7252  
 acatggatca gtgtgtgtcc ccctcaggaa tgtatccaca gtggccttcc ttgctggtgg 7312  
 gcagtgtatc ctgatggcag ggtacaagta ccattaatga agggctctgca acataaagcc 7372  
 ttaaaaagac acacactaag aaaactgtaa aacctgaac attgttattt atatttttta 7432  
 aaatggaaaa gatcactatg ttgtttgtgc taaccactta ttgattctg tttgtggtg 7492  
 gacatagatg attacgtttg agctttgtat ttgtgaaaa ccttaatgaa atgaattcca 7552  
 aagat 7557

&lt;210&gt; 38

&lt;211&gt; 1995

&lt;212&gt; PRT

&lt;213&gt; NM\_020859 ShrmL, Shroom-related protein

&lt;400&gt; 38

Met Arg Thr Thr Glu Asp Phe His Lys Pro Ser Ala Thr Leu Asn Ser  
 1 5 10 15

Asn Thr Ala Thr Lys Gly Arg Tyr Ile Tyr Leu Glu Ala Phe Leu Glu  
 20 25 30

Gly Gly Ala Pro Trp Gly Phe Thr Leu Lys Gly Gly Leu Glu His Gly  
 35 40 45

Glu Pro Leu Ile Ile Ser Lys Val Glu Glu Gly Gly Lys Ala Asp Thr  
 50 55 60

Leu Ser Ser Lys Leu Gln Ala Gly Asp Glu Val Val His Ile Asn Glu  
 65 70 75 80

Val Thr Leu Ser Ser Arg Lys Glu Ala Val Ser Leu Val Lys Gly  
 85 90 95

Ser Tyr Lys Thr Leu Arg Leu Val Val Arg Arg Asp Val Cys Thr Asp  
 100 105 110

Pro Gly His Ala Asp Thr Gly Ala Ser Asn Phe Val Ser Pro Glu His  
 115 120 125

Leu Thr Ser Gly Pro Gln His Arg Lys Ala Ala Trp Ser Gly Gly Val  
130 135 140

Lys Leu Arg Leu Lys His Arg Ser Ser Glu Pro Ala Gly Arg Pro His  
145 150 155 160

Ser Trp His Thr Thr Lys Ser Gly Glu Lys Gln Pro Asp Ala Ser Met  
165 170 175

Met Gln Ile Ser Gln Gly Met Ile Gly Pro Pro Trp His Gln Ser Tyr  
180 185 190

His Ser Ser Ser Ser Thr Ser Asp Leu Ser Asn Tyr Asp His Ala Tyr  
195 200 205

Leu Arg Arg Ser Pro Asp Gln Cys Ser Ser Gln Gly Ser Met Glu Ser  
210 215 220

Leu Glu Pro Ser Gly Ala Tyr Pro Pro Cys His Leu Ser Pro Ala Lys  
225 230 235 240

Ser Thr Gly Ser Ile Asp Gln Leu Ser His Phe His Asn Lys Arg Asp  
245 250 255

Ser Ala Tyr Ser Ser Phe Ser Thr Ser Ser Ser Ile Leu Glu Tyr Pro  
260 265 270

His Pro Gly Ile Ser Ala Arg Glu Arg Ser Gly Ser Met Asp Asn Thr  
275 280 285

Ser Ala Arg Gly Gly Leu Leu Glu Gly Met Arg Gln Ala Asp Ile Arg  
290 295 300

Tyr Val Lys Thr Val Tyr Asp Thr Arg Arg Gly Val Ser Ala Glu Tyr  
305 310 315 320

Glu Val Asn Ser Ser Ala Leu Leu Leu Gln Gly Arg Glu Ala Arg Ala  
325 330 335

Ser Ala Asn Gly Gln Gly Tyr Asp Lys Trp Ser Asn Ile Pro Arg Gly  
340 345 350

Lys Gly Val Pro Pro Pro Ser Trp Ser Gln Gln Cys Pro Ser Ser Leu  
355 360 365

Glu Thr Ala Thr Asp Asn Leu Pro Pro Lys Val Gly Ala Pro Leu Pro  
370 375 380

Pro Ala Arg Ser Asp Ser Tyr Ala Ala Phe Arg His Arg Glu Arg Pro  
385 390 395 400

Ser Ser Trp Ser Ser Leu Asp Gln Lys Arg Leu Cys Arg Pro Gln Ala  
405 410 415

Asn Ser Leu Gly Ser Leu Lys Ser Pro Phe Ile Glu Glu Gln Leu His  
420 425 430

Thr Val Leu Glu Lys Ser Pro Glu Asn Ser Pro Pro Val Lys Pro Lys  
435 440 445

His Asn Tyr Thr Gln Lys Ala Gln Pro Gly Gln Pro Leu Leu Pro Thr  
450 455 460

Ser Ile Tyr Ala Val Pro Ser Leu Glu Pro His Phe Ala Gln Val Pro  
465 470 475 480

Gln Pro Ser Val Ser Ser Asn Gly Met Leu Tyr Pro Ala Leu Ala Lys  
485 490 495

Glu Ser Gly Tyr Ile Ala Pro Gln Gly Ala Cys Asn Lys Met Ala Thr  
500 505 510

Ile Asp Glu Asn Gly Asn Gln Asn Gly Ser Gly Arg Pro Gly Phe Ala  
515 520 525

Phe Cys Gln Pro Leu Glu His Asp Leu Leu Ser Pro Val Glu Lys Lys  
530 535 540

Pro Glu Ala Thr Ala Lys Tyr Val Pro Ser Lys Val His Phe Cys Ser  
545 550 555 560

Val Pro Glu Asn Glu Glu Asp Ala Ser Leu Lys Arg His Leu Thr Pro  
565 570 575

Pro Gln Gly Asn Ser Pro His Ser Asn Glu Arg Lys Ser Thr His Ser  
580 585 590

Asn Lys Pro Ser Ser His Pro His Ser Leu Lys Cys Pro Gln Ala Gln  
595 600 605

Ala Trp Gln Ala Gly Glu Asp Lys Arg Ser Ser Arg Leu Ser Glu Pro  
610 615 620

Trp Glu Gly Asp Phe Gln Glu Asp His Asn Ala Asn Leu Trp Arg Arg  
625 630 635 640

Leu Glu Arg Glu Gly Leu Gly Gln Ser Leu Ser Gly Asn Phe Gly Lys  
645 650 655

Thr Lys Ser Ala Phe Ser Ser Leu Gln Asn Ile Pro Glu Ser Leu Arg  
660 665 670

Arg His Ser Ser Leu Glu Leu Gly Arg Gly Thr Gln Glu Gly Tyr Pro  
675 680 685

Gly Gly Arg Pro Thr Cys Ala Val Asn Thr Lys Ala Glu Asp Pro Gly  
690 695 700

Arg Lys Ala Ala Pro Asp Leu Gly Ser His Leu Asp Arg Gln Val Ser  
705 710 715 720

Tyr Pro Arg Pro Glu Gly Arg Thr Gly Ala Ser Ala Ser Phe Asn Ser  
725 730 735

Thr Asp Pro Ser Pro Glu Glu Pro Pro Ala Pro Ser His Pro His Thr  
740 745 750

Ser Ser Leu Gly Arg Arg Gly Pro Gly Pro Gly Ser Ala Ser Ala Leu  
755 760 765

Gln Gly Phe Gln Tyr Gly Lys Pro His Cys Ser Val Leu Glu Lys Val  
770 775 780

Ser Lys Phe Glu Gln Arg Glu Gln Gly Ser Gln Arg Pro Ser Val Gly  
785 790 795 800

Gly Ser Gly Phe Gly His Asn Tyr Arg Pro His Arg Thr Val Ser Thr  
805 810 815

Ser Ser Thr Ser Gly Asn Asp Phe Glu Glu Thr Lys Ala His Ile Arg  
820 825 830

Phe Ser Glu Ser Ala Glu Pro Leu Gly Asn Gly Glu Gln His Phe Lys  
835 840 845

Asn Gly Glu Leu Lys Leu Glu Glu Ala Ser Arg Gln Pro Cys Gly Gln  
850 855 860

Gln Leu Ser Gly Gly Ala Ser Asp Ser Gly Arg Gly Pro Gln Arg Pro  
865 870 875 880

Asp Ala Arg Leu Leu Arg Ser Gln Ser Thr Phe Gln Leu Ser Ser Glu  
885 890 895

Pro Glu Arg Glu Pro Glu Trp Arg Asp Arg Pro Gly Ser Pro Glu Ser  
900 905 910

Pro Leu Leu Asp Ala Pro Phe Ser Arg Ala Tyr Arg Asn Ser Ile Lys  
915 920 925

Asp Ala Gln Ser Arg Val Leu Gly Ala Thr Ser Phe Arg Arg Arg Asp  
930 935 940

Leu Glu Leu Gly Ala Pro Val Ala Ser Arg Ser Trp Arg Pro Arg Pro  
945 950 955 960

Ser Ser Ala His Val Gly Leu Arg Ser Pro Glu Ala Ser Ala Ser Ala  
965 970 975

Ser Pro His Thr Pro Arg Glu Arg His Ser Val Thr Pro Ala Glu Gly  
980 985 990

Asp Leu Ala Arg Pro Val Pro Pro Ala Ala Arg Arg Gly Ala Arg Arg  
995 1000 1005

Arg Leu Thr Pro Glu Gln Lys Lys Arg Ser Tyr Ser Glu Pro Glu  
1010 1015 1020

Lys Met Asn Glu Val Gly Ile Val Glu Glu Ala Glu Pro Ala Pro  
1025 1030 1035

Leu Gly Pro Gln Arg Asn Gly Met Arg Phe Pro Glu Ser Ser Val  
1040 1045 1050

Ala Asp Arg Arg Arg Leu Phe Glu Arg Asp Gly Lys Ala Cys Ser  
1055 1060 1065

Thr Leu Ser Leu Ser Gly Pro Glu Leu Lys Gln Phe Gln Gln Ser  
1070 1075 1080

Ala Leu Ala Asp Tyr Ile Gln Arg Lys Thr Gly Lys Arg Pro Thr  
1085 1090 1095

Ser Ala Ala Gly Cys Ser Leu Gln Glu Pro Gly Pro Leu Arg Glu  
1100 1105 1110

Arg Ala Gln Ser Ala Tyr Leu Gln Pro Gly Pro Ala Ala Leu Glu  
1115 1120 1125

Gly Ser Gly Leu Ala Ser Ala Ser Ser Leu Ser Ser Leu Arg Glu  
1130 1135 1140

Pro Ser Leu Gln Pro Arg Arg Glu Ala Thr Leu Leu Pro Ala Thr  
1145 1150 1155

Val Ala Glu Thr Gln Gln Ala Pro Arg Asp Arg Ser Ser Ser Phe  
1160 1165 1170

Ala Gly Gly Arg Arg Leu Gly Glu Arg Arg Arg Gly Asp Leu Leu  
1175 1180 1185

Ser Gly Ala Asn Gly Gly Thr Arg Gly Thr Gln Arg Gly Asp Glu  
1190 1195 1200

Thr Pro Arg Glu Pro Ser Ser Trp Gly Ala Arg Ala Gly Lys Ser  
 1205 1210 1215  
 Met Ser Ala Glu Asp Leu Leu Glu Arg Ser Asp Val Leu Ala Gly  
 1220 1225 1230  
 Pro Val His Val Arg Ser Arg Ser Ser Pro Ala Thr Ala Asp Lys  
 1235 1240 1245  
 Arg Gln Asp Val Leu Leu Gly Gln Asp Ser Gly Phe Gly Leu Val  
 1250 1255 1260  
 Lys Asp Pro Cys Tyr Leu Ala Gly Pro Gly Ser Arg Ser Leu Ser  
 1265 1270 1275  
 Cys Ser Glu Arg Gly Gln Glu Glu Met Leu Leu Leu Phe His His  
 1280 1285 1290  
 Leu Thr Pro Arg Trp Gly Gly Ser Gly Cys Lys Ala Ile Gly Asp  
 1295 1300 1305  
 Ser Ser Val Pro Ser Glu Cys Pro Gly Thr Leu Asp His Gln Arg  
 1310 1315 1320  
 Gln Ala Ser Arg Thr Pro Cys Pro Arg Pro Pro Leu Ala Gly Thr  
 1325 1330 1335  
 Gln Gly Leu Val Thr Asp Thr Arg Ala Ala Pro Leu Thr Pro Ile  
 1340 1345 1350  
 Gly Thr Pro Leu Pro Ser Ala Ile Pro Ser Gly Tyr Cys Ser Gln  
 1355 1360 1365  
 Asp Gly Gln Thr Gly Arg Gln Pro Leu Pro Pro Tyr Thr Pro Ala  
 1370 1375 1380  
 Met Met His Arg Ser Asn Gly His Thr Leu Thr Gln Pro Pro Gly  
 1385 1390 1395  
 Pro Arg Gly Cys Glu Gly Asp Gly Pro Glu His Gly Val Glu Glu  
 1400 1405 1410  
 Gly Thr Arg Lys Arg Val Ser Leu Pro Gln Trp Pro Pro Pro Ser  
 1415 1420 1425  
 Arg Ala Lys Trp Ala His Ala Ala Arg Glu Asp Ser Leu Pro Glu  
 1430 1435 1440  
 Glu Ser Ser Ala Pro Asp Phe Ala Asn Leu Lys His Tyr Gln Lys  
 1445 1450 1455



Gln Gln Ser Leu Pro Ser Leu Cys Ser Thr Ser Asp Pro Asp Thr  
1460 1465 1470

Pro Leu Gly Ala Pro Ser Thr Pro Gly Arg Ile Ser Leu Arg Ile  
1475 1480 1485

Ser Glu Ser Val Leu Arg Asp Ser Pro Pro Pro His Glu Asp Tyr  
1490 1495 1500

Glu Asp Glu Val Phe Val Arg Asp Pro His Pro Lys Ala Thr Ser  
1505 1510 1515

Ser Pro Thr Phe Glu Pro Leu Pro Pro Pro Pro Pro Pro Pro Pro  
1520 1525 1530

Ser Gln Glu Thr Pro Val Tyr Ser Met Asp Asp Phe Pro Pro Pro  
1535 1540 1545

Pro Pro His Thr Val Cys Glu Ala Gln Leu Asp Ser Glu Asp Pro  
1550 1555 1560

Glu Gly Pro Arg Pro Ser Phe Asn Lys Leu Ser Lys Val Thr Ile  
1565 1570 1575

Ala Arg Glu Arg His Met Pro Gly Ala Ala His Val Val Gly Ser  
1580 1585 1590

Gln Thr Leu Ala Ser Arg Leu Gln Thr Ser Ile Lys Gly Ser Glu  
1595 1600 1605

Ala Glu Ser Thr Pro Pro Ser Phe Met Ser Val His Ala Gln Leu  
1610 1615 1620

Ala Gly Ser Leu Gly Gly Gln Pro Ala Pro Ile Gln Thr Gln Ser  
1625 1630 1635

Leu Ser His Asp Pro Val Ser Gly Thr Gln Gly Leu Glu Lys Lys  
1640 1645 1650

Val Ser Pro Asp Pro Gln Lys Ser Ser Glu Asp Ile Arg Thr Glu  
1655 1660 1665

Ala Leu Ala Lys Glu Ile Val His Gln Asp Lys Ser Leu Ala Asp  
1670 1675 1680

Ile Leu Asp Pro Asp Ser Arg Leu Lys Thr Thr Met Asp Leu Met  
1685 1690 1695

Glu Gly Leu Phe Pro Arg Asp Val Asn Leu Leu Lys Glu Asn Ser  
1700 1705 1710

Val Lys Arg Lys Ala Ile Gln Arg Thr Val Ser Ser Ser Gly Cys  
1715 1720 1725

Glu Gly Lys Arg Asn Glu Asp Lys Glu Ala Val Ser Met Leu Val  
1730 1735 1740

Asn Cys Pro Ala Tyr Tyr Ser Val Ser Ala Pro Lys Ala Glu Leu  
1745 1750 1755

Leu Asn Lys Ile Lys Glu Met Pro Ala Glu Val Asn Glu Glu Glu  
1760 1765 1770

Glu Gln Ala Asp Val Asn Glu Lys Lys Ala Glu Leu Ile Gly Ser  
1775 1780 1785

Leu Thr His Lys Leu Glu Thr Leu Gln Glu Ala Lys Gly Ser Leu  
1790 1795 1800

Leu Thr Asp Ile Lys Leu Asn Asn Ala Leu Gly Glu Glu Val Glu  
1805 1810 1815

Ala Leu Ile Ser Glu Leu Cys Lys Pro Asn Glu Phe Asp Lys Tyr  
1820 1825 1830

Arg Met Phe Ile Gly Asp Leu Asp Lys Val Val Asn Leu Leu Leu  
1835 1840 1845

Ser Leu Ser Gly Arg Leu Ala Arg Val Glu Asn Val Leu Ser Gly  
1850 1855 1860

Leu Gly Glu Asp Ala Ser Asn Glu Glu Arg Ser Ser Leu Tyr Glu  
1865 1870 1875

Lys Arg Lys Ile Leu Ala Gly Gln His Glu Asp Ala Arg Glu Leu  
1880 1885 1890

Lys Glu Asn Leu Asp Arg Arg Glu Arg Val Val Leu Gly Ile Leu  
1895 1900 1905

Ala Asn Tyr Leu Ser Glu Glu Gln Leu Gln Asp Tyr Gln His Phe  
1910 1915 1920

Val Lys Met Lys Ser Thr Leu Leu Ile Glu Gln Arg Lys Leu Asp  
1925 1930 1935

Asp Lys Ile Lys Leu Gly Gln Glu Gln Val Lys Cys Leu Leu Glu  
1940 1945 1950

Ser Leu Pro Ser Asp Phe Ile Pro Lys Ala Gly Ala Leu Ala Leu  
1955 1960 1965

Pro Pro Asn Leu Thr Ser Glu Pro Ile Pro Ala Gly Gly Cys Thr  
1970 1975 1980

Phe Ser Gly Ile Phe Pro Thr Leu Thr Ser Pro Leu  
1985 1990 1995

<210> 39

<211> 2529

<212> DNA

<213> NM\_004433 ELF3, E74-like factor 3

<220>

<221> CDS

<222> (203)..(1315)

<223>

<400> 39  
gagagagaga gagagagaga actagtctcg aggtcgacgc ggccgcgaat tcgcggccgc 60  
gtcgaccgaa gacgcaggct ctatttagag ccgggtaggg gagcgacagg ccagatacct 120  
cagcgctacc tggcggaact ggatttctct cccgcctgcc ggctgcctg ccacagccgg 180  
actccgccac tccggtagcc tc atg gct gca acc tgt gag att agc aac att 232  
Met Ala Ala Thr Cys Glu Ile Ser Asn Ile  
1 5 10  
ttt agc aac tac ttc agt gcg atg tac agc tcg gag gac tcc acc ctg 280  
Phe Ser Asn Tyr Phe Ser Ala Met Tyr Ser Ser Glu Asp Ser Thr Leu  
15 20 25  
gcc tct gtt ccc cct gct gcc acc ttt ggg gcc gat gac ttg gta ctg 328  
Ala Ser Val Pro Pro Ala Ala Thr Phe Gly Ala Asp Asp Leu Val Leu  
30 35 40  
acc ctg agc aac ccc cag atg tca ttg gag ggt aca gag aag gcc agc 376  
Thr Leu Ser Asn Pro Gln Met Ser Leu Glu Gly Thr Glu Lys Ala Ser  
45 50 55  
tgg ttg ggg gaa cag ccc cag ttc tgg tcg aag acg cag gtt ctg gac 424  
Trp Leu Gly Glu Gln Pro Gln Phe Trp Ser Lys Thr Gln Val Leu Asp  
60 65 70  
tgg atc agc tac caa gtg gag aag aac aag tac gac gca agc gcc att 472  
Trp Ile Ser Tyr Gln Val Glu Lys Asn Lys Tyr Asp Ala Ser Ala Ile  
75 80 85 90  
gac ttc tca cga tgt gac atg gat ggc gcc acc ctc tgc aat tgt gcc 520  
Asp Phe Ser Arg Cys Asp Met Asp Gly Ala Thr Leu Cys Asn Cys Ala  
95 100 105  
ctt gag gag ctg cgt ctg gtc ttt ggg cct ctg.ggg gac caa ctc cat 568  
Leu Glu Glu Leu Arg Leu Val Phe Gly Pro Leu Gly Asp Gln Leu His  
110 115 120

|   |      |
|---|------|
| gcc cag ctg cga gac ctc act tcc agc tct tct gat gag ctc agt tgg<br>Ala Gln Leu Arg Asp Leu Thr Ser Ser Ser Ser Asp Glu Leu Ser Trp<br>125 130 135     | 616  |
| atc att gag ctg ctg gag aag gat ggc atg gcc ttc cag gag gcc cta<br>Ile Ile Glu Leu Leu Glu Lys Asp Gly Met Ala Phe Gln Glu Ala Leu<br>140 145 150     | 664  |
| gac cca ggg ccc ttt gac cag ggc agc ccc ttt gcc cag gag ctg ctg<br>Asp Pro Gly Pro Phe Asp Gln Gly Ser Pro Phe Ala Gln Glu Leu Leu<br>155 160 165 170 | 712  |
| gac gac ggt cag caa gcc agc ccc tac caa ccc ggc agc tgt ggc gca<br>Asp Asp Gly Gln Gln Ala Ser Pro Tyr His Pro Gly Ser Cys Gly Ala<br>175 180 185     | 760  |
| gga gcc ccc tcc ccc ggc agc tct gac gtc tcc acc gca ggg act ggt<br>Gly Ala Pro Ser Pro Gly Ser Ser Asp Val Ser Thr Ala Gly Thr Gly<br>190 195 200     | 808  |
| gct tct cgg agc tcc cac tcc tca gac tcc ggt gga agt gac gtg gac<br>Ala Ser Arg Ser Ser His Ser Ser Asp Ser Gly Gly Ser Asp Val Asp<br>205 210 215     | 856  |
| ctg gat ccc act gat ggc aag ctc ttc ccc agc gat ggt ttt cgt gac<br>Leu Asp Pro Thr Asp Gly Lys Leu Phe Pro Ser Asp Gly Phe Arg Asp<br>220 225 230     | 904  |
| tgc aag aag ggg gat ccc aag cac ggg aag cgg aaa cga ggc cgg ccc<br>Cys Lys Lys Gly Asp Pro Lys His Gly Lys Arg Lys Arg Gly Arg Pro<br>235 240 245 250 | 952  |
| cga aag ctg agc aaa gag tac tgg gac tgt ctc gag ggc aag aag agc<br>Arg Lys Leu Ser Lys Glu Tyr Trp Asp Cys Leu Glu Gly Lys Lys Ser<br>255 260 265     | 1000 |
| aag cac gcg ccc aga ggc acc cac ctg tgg gag ttc atc cgg gac atc<br>Lys His Ala Pro Arg Gly Thr His Leu Trp Glu Phe Ile Arg Asp Ile<br>270 275 280     | 1048 |
| ctc atc cac ccg gag ctc aac gag ggc ctc atg aag tgg gag aat cgg<br>Leu Ile His Pro Glu Leu Asn Glu Gly Leu Met Lys Trp Glu Asn Arg<br>285 290 295     | 1096 |
| cat gaa ggc gtc ttc aag ttc ctg cgc tcc gag gct gtg gcc caa cta<br>His Glu Gly Val Phe Lys Phe Leu Arg Ser Glu Ala Val Ala Gln Leu<br>300 305 310     | 1144 |
| tgg ggc caa aag aaa aag aac agc aac atg acc tac gag aag ctg agc<br>Trp Gly Gln Lys Lys Lys Asn Ser Asn Met Thr Tyr Glu Lys Leu Ser<br>315 320 325 330 | 1192 |
| cgg gcc atg agg tac tac tac aaa cgg gag atc ctg gaa cgg gtg gat<br>Arg Ala Met Arg Tyr Tyr Tyr Lys Arg Glu Ile Leu Glu Arg Val Asp<br>335 340 345     | 1240 |
| ggc cgg cga ctc gtc tac aag ttt ggc aaa aac tca agc ggc tgg aag<br>Gly Arg Arg Leu Val Tyr Lys Phe Gly Lys Asn Ser Ser Gly Trp Lys<br>350 355 360     | 1288 |
| gag gaa gag gtt ctc cag agt cgg aac tgagggttg aactataccc<br>Glu Glu Glu Val Leu Gln Ser Arg Asn<br>365 370  | 1335 |
| gggaccaaac tcacggacca ctcgaggcct gcaaaccttc ctgggaggac aggcaggcca   | 1395 |
| gatggcccct ccactgggga atgctcccag ctgtgctgtg gagagaagct gatgttttgg   | 1455 |

tgtattgtca gccatcgctc ttggactcgg agactatggc ctgcctccc caccctectc 1515  
 ttggaattac aagccctggg gtttgaagct gactttatag ctgcaagtgt atctcctttt 1575  
 atctgggtgcc tcctcaaacc cagtctcaga cactaaatgc agacaacacc ttcctcctgc 1635  
 agacacttgg actgagccaa ggaggcttgg gaggccctag ggagcaccgt gatggagagg 1695  
 acagagcagg ggtccagca cttctttctg gactggcggt cacctccctg ctgagtgttt 1755  
 gggctccacg ggcaggggtc agagcactcc ctaatttatg tgctatataa atatgtcaga 1815  
 tgtacataga gatctatttt ttctaaaaca ttcccctccc cactcctctc ccacagagtg 1875  
 ctggactgtt ccaggccctc cagtgggctg atgctgggac ccttaggatg gggctcccag 1935  
 ctctctttct ctgtgaatgg aggcagagac ctccaataaa gtgccttctg ggctttttct 1995  
 aacctttgtc ttagctacct gtgtactgaa atttgggcct ttggatcgaa tatggtcaag 2055  
 aggttggagg ggaggaaaat gaaggtctac caggatgagg gtgagggcaa aggctgacga 2115  
 agaggggagt tacagatttc ctgtagcagg tgtgggctta cagacacatg gactgggctg 2175  
 ggaggcgagc aaaggaagca gctgagactg ttggagaacg ttacaagact tcatgcaaga 2235  
 aggacatgaa ctgagaacac tgaggtcaga agcatctgct gtcacgacac cgctcgagtg 2295  
 accttgacct tgaccaagtc ttgtccttgt ttaggactga tttttcctat taggctaggg 2355  
 tttggacctg atgttctcaa gatgtctaga attgcatggc tggccttctg gaatagatgg 2415  
 ttttgcatc cagccaagtg tgctgtaaac tgtatatctg taatatgaat ccagctttt 2475  
 gagtctgaca aatcagagt taggatcttg taaaggaaaa aaaaaaaaaa aaaa 2529

<210> 40

<211> 371

<212> PRT

<213> NM\_004433 ELF3, E74-like factor 3

<400> 40

Met Ala Ala Thr Cys Glu Ile Ser Asn Ile Phe Ser Asn Tyr Phe Ser  
1 5 10 15

Ala Met Tyr Ser Ser Glu Asp Ser Thr Leu Ala Ser Val Pro Pro Ala  
20 25 30

Ala Thr Phe Gly Ala Asp Asp Leu Val Leu Thr Leu Ser Asn Pro Gln  
35 40 45

Met Ser Leu Glu Gly Thr Glu Lys Ala Ser Trp Leu Gly Glu Gln Pro  
50 55 60

Gln Phe Trp Ser Lys Thr Gln Val Leu Asp Trp Ile Ser Tyr Gln Val  
65 70 75 80

Glu Lys Asn Lys Tyr Asp Ala Ser Ala Ile Asp Phe Ser Arg Cys Asp  
85 90 95

Met Asp Gly Ala Thr Leu Cys Asn Cys Ala Leu Glu Glu Leu Arg Leu  
100 105 110

Val Phe Gly Pro Leu Gly Asp Gln Leu His Ala Gln Leu Arg Asp Leu  
115 120 125

Thr Ser Ser Ser Ser Asp Glu Leu Ser Trp Ile Ile Glu Leu Leu Glu  
130 135 140

Lys Asp Gly Met Ala Phe Gln Glu Ala Leu Asp Pro Gly Pro Phe Asp  
145 150 155 160

Gln Gly Ser Pro Phe Ala Gln Glu Leu Leu Asp Asp Gly Gln Gln Ala  
165 170 175

Ser Pro Tyr His Pro Gly Ser Cys Gly Ala Gly Ala Pro Ser Pro Gly  
180 185 190

Ser Ser Asp Val Ser Thr Ala Gly Thr Gly Ala Ser Arg Ser Ser His  
195 200 205

Ser Ser Asp Ser Gly Gly Ser Asp Val Asp Leu Asp Pro Thr Asp Gly  
210 215 220

Lys Leu Phe Pro Ser Asp Gly Phe Arg Asp Cys Lys Lys Gly Asp Pro  
225 230 235 240

Lys His Gly Lys Arg Lys Arg Gly Arg Pro Arg Lys Leu Ser Lys Glu  
245 250 255

Tyr Trp Asp Cys Leu Glu Gly Lys Lys Ser Lys His Ala Pro Arg Gly  
260 265 270

Thr His Leu Trp Glu Phe Ile Arg Asp Ile Leu Ile His Pro Glu Leu  
275 280 285

Asn Glu Gly Leu Met Lys Trp Glu Asn Arg His Glu Gly Val Phe Lys  
290 295 300

Phe Leu Arg Ser Glu Ala Val Ala Gln Leu Trp Gly Gln Lys Lys Lys  
305 310 315 320

Asn Ser Asn Met Thr Tyr Glu Lys Leu Ser Arg Ala Met Arg Tyr Tyr  
325 330 335

Tyr Lys Arg Glu Ile Leu Glu Arg Val Asp Gly Arg Arg Leu Val Tyr

340

345

350

Lys Phe Gly Lys Asn Ser Ser Gly Trp Lys Glu Glu Glu Val Leu Gln  
 355 360 365

Ser Arg Asn  
 370

&lt;210&gt; 41

&lt;211&gt; 4020

&lt;212&gt; DNA

&lt;213&gt; NM\_019027 FLJ20273

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (240)..(1811)

&lt;223&gt;

&lt;400&gt; 41

gtcaattgcc ctttagtccc aggactaacc ggaagcttct gcaacaggag gacattgaaa 60

ataagatgga acccatccac ataaggattt gcctcaaagg gcaactgcaaa aattgaacag 120

aggaatccca aggaagctgc ctgaatttgc ctgtatactc tcgtttctgcg acttataaag 180

gaccagacaa atcaaattag tgggttttgggt ttccgccagc tgtggatgcc tttgacatt 239

atg acc gca gag gat tcc acc gca gcc atg agc agt gac tcg gcc gcc 287  
 Met Thr Ala Glu Asp Ser Thr Ala Ala Met Ser Ser Asp Ser Ala Ala  
 1 5 10 15

ggg tcc tcg gcc aag gtg ccc gag ggc gtg gcg ggc gcg ccc aac gag 335  
 Gly Ser Ser Ala Lys Val Pro Glu Gly Val Ala Gly Ala Pro Asn Glu  
 20 25 30

gca gca ctg ctg gcg ctg atg gag cgc acg ggc tac agc atg gtg caa 383  
 Ala Ala Leu Leu Ala Leu Met Glu Arg Thr Gly Tyr Ser Met Val Gln  
 35 40 45

gag aac ggg cag cgc aag tac ggc ggc cca ccg ccc ggc tgg gag ggc 431  
 Glu Asn Gly Gln Arg Lys Tyr Gly Gly Pro Pro Pro Gly Trp Glu Gly  
 50 55 60

ccg cac ccg cag cgt ggc tgc gag gtc ttc gtg ggc aag atc ccg cgc 479  
 Pro His Pro Gln Arg Gly Cys Glu Val Phe Val Gly Lys Ile Pro Arg  
 65 70 75 80

gac gtg tac gag gac gag ctg gtg ccc gtg ttc gag gcc gtg ggc cgc 527  
 Asp Val Tyr Glu Asp Glu Leu Val Pro Val Phe Glu Ala Val Gly Arg  
 85 90 95

acc tac gag ctg cgc ctc atg atg gac ttt gac ggc aag aac cgc ggc 575  
 Thr Tyr Glu Leu Arg Leu Met Met Asp Phe Asp Gly Lys Asn Arg Gly  
 100 105 110

|   |      |
|---|------|
| tac gcc ttc gtc atg tac tgc cac aag cac gag gcc aag cgc gca gtg<br>Tyr Ala Phe Val Met Tyr Cys His Lys His Glu Ala Lys Arg Ala Val<br>115 120 125     | 623  |
| cgt gag ctc aac aac tac gag atc cgc ccg gcc cgc ctg ctc gcc gtg<br>Arg Glu Leu Asn Asn Tyr Glu Ile Arg Pro Gly Arg Leu Leu Gly Val<br>130 135 140     | 671  |
| tgc tgc agc gtg gac aac tgc cgc ctc ttc atc gcc ggc atc ccc aag<br>Cys Cys Ser Val Asp Asn Cys Arg Leu Phe Ile Gly Gly Ile Pro Lys<br>145 150 155 160 | 719  |
| atg aag aag cgc gag gaa atc ctg gag gag att gcc aag gtc acc gag<br>Met Lys Lys Arg Glu Glu Ile Leu Glu Ile Ala Lys Val Thr Glu<br>165 170 175         | 767  |
| ggc gtg ctg gac gtg atc gtc tac gcc agc gcg gcc gac aag atg aag<br>Gly Val Leu Asp Val Ile Val Tyr Ala Ser Ala Ala Asp Lys Met Lys<br>180 185 190     | 815  |
| aac cgc gcc ttc gcc ttc gtg gag tac gag agc cac cgc gcg gct gcc<br>Asn Arg Gly Phe Ala Phe Val Glu Tyr Glu Ser His Arg Ala Ala Ala<br>195 200 205     | 863  |
| atg gct cgc cgc aag ctc atg cct gcc cgc atc cag ctg tgg gcc cac<br>Met Ala Arg Arg Lys Leu Met Pro Gly Arg Ile Gln Leu Trp Gly His<br>210 215 220     | 911  |
| cag atc gcc gtg gac tgg gcc gag cct gag atc gac gtg gac gag gac<br>Gln Ile Ala Val Asp Trp Ala Glu Pro Glu Ile Asp Val Asp Glu Asp<br>225 230 235 240 | 959  |
| gtg atg gag acc gtg aag atc ctc tac gtg cgc aac ctc atg atc gag<br>Val Met Glu Thr Val Lys Ile Leu Tyr Val Arg Asn Leu Met Ile Glu<br>245 250 255     | 1007 |
| acc acc gag gac acc atc aag aag agc ttc gcc cag ttc aac ccc gcc<br>Thr Thr Glu Asp Thr Ile Lys Lys Ser Phe Gly Gln Phe Asn Pro Gly<br>260 265 270     | 1055 |
| tgc gtg gag cgc gtc aag aag atc cgc gac tac gcc ttc gtg cac ttc<br>Cys Val Glu Arg Val Lys Lys Ile Arg Asp Tyr Ala Phe Val His Phe<br>275 280 285     | 1103 |
| acc agc cgc gag gat gcc gtg cat gcc atg aac aac ctc aac gcc act<br>Thr Ser Arg Glu Asp Ala Val His Ala Met Asn Asn Leu Asn Gly Thr<br>290 295 300     | 1151 |
| gag ctg gag gcc tcg tgc ctg gag gtc acg ctg gcc aag ccc gtg gac<br>Glu Leu Glu Gly Ser Cys Leu Glu Val Thr Leu Ala Lys Pro Val Asp<br>305 310 315 320 | 1199 |
| aag gag cag tac tcg cgc tac cag aag gca gcc agg gcc gcc gcc gcg<br>Lys Glu Gln Tyr Ser Arg Tyr Gln Lys Ala Ala Arg Gly Gly Gly Ala<br>325 330 335     | 1247 |
| gct gag gca gcg cag cag ccc agc tac gtg tac tcc tgc gac ccc tac<br>Ala Glu Ala Ala Gln Gln Pro Ser Tyr Val Tyr Ser Cys Asp Pro Tyr<br>340 345 350     | 1295 |
| aca ctg gcc tac tac gcc tac ccc tac aac gcg ctc att ggg ccc aac<br>Thr Leu Ala Tyr Tyr Gly Tyr Pro Tyr Asn Ala Leu Ile Gly Pro Asn<br>355 360 365     | 1343 |
| agg gac tac ttt gtg aaa gta gcc atc cct gcc att ggg gct cag tat<br>Arg Asp Tyr Phe Val Lys Val Ala Ile Pro Ala Ile Gly Ala Gln Tyr<br>370 375 380     | 1391 |



|   |      |
|---|------|
| tcc atg ttt cca gca gct cca gcc cct aaa atg att gaa gat ggc aaa<br>Ser Met Phe Pro Ala Ala Pro Ala Pro Lys Met Ile Glu Asp Gly Lys<br>385 390 395 400 | 1439 |
| atc cac aca gtg gag cac atg atc agc ccc att gct gtg cag cca gac<br>Ile His Thr Val Glu His Met Ile Ser Pro Ile Ala Val Gln Pro Asp<br>405 410 415     | 1487 |
| cca gcc agt gct gct gcc gcc gca gcc gcg gcc gca gcc gcc gca gcc<br>Pro Ala Ser Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala<br>420 425 430         | 1535 |
| gct gtc att ccc act gtg tgg acg cca cca cct ttc cag ggc cgc cca<br>Ala Val Ile Pro Thr Val Ser Thr Pro Pro Pro Phe Gln Gly Arg Pro<br>435 440 445     | 1583 |
| ata act cca gta tac acg gtg gct cca aac gtt cag aga att cct act<br>Ile Thr Pro Val Tyr Thr Val Ala Pro Asn Val Gln Arg Ile Pro Thr<br>450 455 460     | 1631 |
| gcc ggg atc tac ggg gcc agt tac gtg cca ttt gct gct cca gct aca<br>Ala Gly Ile Tyr Gly Ala Ser Tyr Val Pro Phe Ala Ala Pro Ala Thr<br>465 470 475 480 | 1679 |
| gcc acg atc gcc aca cta cag aag aac gcg gca gcc gcg gcc gcc gtg<br>Ala Thr Ile Ala Thr Leu Gln Lys Asn Ala Ala Ala Ala Ala Val<br>485 490 495         | 1727 |
| tat gga gga tac gca ggc tac ata cct cag gcc ttc cct gct gct gcc<br>Tyr Gly Gly Tyr Ala Gly Tyr Ile Pro Gln Ala Phe Pro Ala Ala Ala<br>500 505 510     | 1775 |
| att cag gtc ccc atc ccc gac gtc tac cag aca tac tgaggctggt<br>Ile Gln Val Pro Ile Pro Asp Val Tyr Gln Thr Tyr<br>515 520                              | 1821 |
| gaccagcagc aagacagacc acacaaacac cactgaagga acgcttgact atttatgaag   | 1881 |
| aaggaacatg ttggattcac acatgcaacc tgaaagtga gaattgttagc agattttatt   | 1941 |
| ctgaattatt ttatatacat gaagttttca ctagtgtttt aagactattt tcaacttagc   | 2001 |
| atgcctacgt tcatacatTTT ccaaaagact tgcaatgggt cgtgccttca ttccatcttt  | 2061 |
| taaaaatttg tatgctgtac tacatttgta tagaggtttt tgttgttgtt tttttaagga   | 2121 |
| tatatTTTtca gtatgaaggt tattttctta acttctgcac tccagagatt tctattttgt  | 2181 |
| agtaccttca ataatatatc aactatatat taaaaaagca cacttgagga gctagggaac   | 2241 |
| tattttgaaa aatatataca atatttaaag ataacaacag tagtgcttaa aaatactaca   | 2301 |
| taaagcatta ttttaaaggt tatactggaa agtgcaattt taaaatgagt aaaacctctg   | 2361 |
| tattttctgct ggcattaagg gttgatgggt ttaccatgta tcatcatggc ggtactattt  | 2421 |
| tttaaaagaa attaaacact ggatctctcc ttaagccaac attgaaaaga cttgccgcac   | 2481 |
| ttctgagtc aaacactgga aagctctcct tgccaccgtt agccggggct cattctccat  | 2541 |
| gtgccttagc cttaaacaatg cccccactcc cacatctctc accctgtccc ctctcccca   | 2601 |
| gattcccaat cccaccgcaa tgtttgcaa gcctaggact gataagtagc tctgatagag  | 2661 |
| gagctggtgg cttttatact tcttctctgg tttttgttgg ggtttgttgt ttogttgttt   | 2721 |
| tttgTTTTtt ttttgtttgg ttggggaagt attgtctctc acgtgtgcta ttttcagtag   | 2781 |

cagagtaagc acaaggtttt aatcgagttg cataagacac ctttgcatag ctattttaatt 2841  
 gcccaatgta aaactttaat gccatttcta atgcttttat tcatttttga agtatgagtt 2901  
 tgtagggaca aagaatgtat gttatcgtag acaagacccc cagagactct tttcagcaga 2961  
 aagttatgct tctagttgcc ttacatggtt tcttgcaaaa ctgtccatgg tcctcaaggg 3021  
 tgttggaac attatgttta ttaaatgggc ctctcttctt ttgctgtgca cttgatgggt 3081  
 gaactggatt ggggtgtgca catccaggag gaggaggaga gacctgtaga agtttaaaga 3141  
 tagtttgtaa atatcttcta atgcttgttt ttagtccttt tatgttgag aagttcatgg 3201  
 tatgtagttt aatgcaaat gaaaccattt tatttcaatg ttattaaaaa ggtttgtttt 3261  
 attaggaagt taatgtattg ttgcagtgtt ttgtgcctgt ttaaaggctt ttgtttagca 3321  
 gagtgaatgt aaaatacagt aaaatgttaa gattgtcatc tactttttta aaaaaaatat 3381  
 caacttgaa ttgtttttta aaggctcaat caaggaagtg aggtgtgcaa taaggtagca 3441  
 agtaaacgc agttgcgttt ttatgtcatg ttagagatcc atacaatttt ccactcacgg 3501  
 gattttgtt gatgctgaa ttcttgtgga ttcataagag gatcatgcc ttagcaagta 3561  
 cttttgtttt gttttaaatt aagagattcc caaatgcctt tttccccctc atcttgaaat 3621  
 gagatgagtt tttatgtgta agcaatattt atttaactat totataaaat tattgagtgc 3681  
 ctactgaggc ctttaagcac cgctaacatt cctttccatc attcttttga atgacataaa 3741  
 ataattgtgc aatgttctgt atgatgtacc ccacagctgc attcaaactc aaatctgtgg 3801  
 gaatgagtga ctcgaccaa tgtaattcgg atcagatcct catccctga ctgtgtgaaa 3861  
 aaagtactct ccttctagt gaggattgtc acagagtttc actggatgaa actatgacct 3921  
 agtattctta ctgtatttta catgtgcctg taaattattt tgccgaaata agaagaagaa 3981  
 gaggaagaaa gaacagtaga aaaaaaaaaa aaaaaaaaaa 4020

&lt;210&gt; 42

&lt;211&gt; 524

&lt;212&gt; PRT

&lt;213&gt; NM\_019027 FLJ20273

&lt;400&gt; 42

Met Thr Ala Glu Asp Ser Thr Ala Ala Met Ser Ser Asp Ser Ala Ala  
 1 5 10 15

Gly Ser Ser Ala Lys Val Pro Glu Gly Val Ala Gly Ala Pro Asn Glu  
 20 25 30

Ala Ala Leu Leu Ala Leu Met Glu Arg Thr Gly Tyr Ser Met Val Gln  
 35 40 45

Glu Asn Gly Gln Arg Lys Tyr Gly Gly Pro Pro Pro Gly Trp Glu Gly

50 55 60

Pro His Pro Gln Arg Gly Cys Glu Val Phe Val Gly Lys Ile Pro Arg  
65 70 75 80

Asp Val Tyr Glu Asp Glu Leu Val Pro Val Phe Glu Ala Val Gly Arg  
85 90 95

Thr Tyr Glu Leu Arg Leu Met Met Asp Phe Asp Gly Lys Asn Arg Gly  
100 105 110

Tyr Ala Phe Val Met Tyr Cys His Lys His Glu Ala Lys Arg Ala Val  
115 120 125

Arg Glu Leu Asn Asn Tyr Glu Ile Arg Pro Gly Arg Leu Leu Gly Val  
130 135 140

Cys Cys Ser Val Asp Asn Cys Arg Leu Phe Ile Gly Gly Ile Pro Lys  
145 150 155 160

Met Lys Lys Arg Glu Glu Ile Leu Glu Glu Ile Ala Lys Val Thr Glu  
165 170 175

Gly Val Leu Asp Val Ile Val Tyr Ala Ser Ala Ala Asp Lys Met Lys  
180 185 190

Asn Arg Gly Phe Ala Phe Val Glu Tyr Glu Ser His Arg Ala Ala Ala  
195 200 205

Met Ala Arg Arg Lys Leu Met Pro Gly Arg Ile Gln Leu Trp Gly His  
210 215 220

Gln Ile Ala Val Asp Trp Ala Glu Pro Glu Ile Asp Val Asp Glu Asp  
225 230 235 240

Val Met Glu Thr Val Lys Ile Leu Tyr Val Arg Asn Leu Met Ile Glu  
245 250 255

Thr Thr Glu Asp Thr Ile Lys Lys Ser Phe Gly Gln Phe Asn Pro Gly  
260 265 270

Cys Val Glu Arg Val Lys Lys Ile Arg Asp Tyr Ala Phe Val His Phe  
275 280 285

Thr Ser Arg Glu Asp Ala Val His Ala Met Asn Asn Leu Asn Gly Thr  
290 295 300

Glu Leu Glu Gly Ser Cys Leu Glu Val Thr Leu Ala Lys Pro Val Asp  
305 310 315 320

Lys Glu Gln Tyr Ser Arg Tyr Gln Lys Ala Ala Arg Gly Gly Gly Ala

325

330

335

Ala Glu Ala Ala Gln Gln Pro Ser Tyr Val Tyr Ser Cys Asp Pro Tyr  
340 345 350

Thr Leu Ala Tyr Tyr Gly Tyr Pro Tyr Asn Ala Leu Ile Gly Pro Asn  
355 360 365

Arg Asp Tyr Phe Val Lys Val Ala Ile Pro Ala Ile Gly Ala Gln Tyr  
370 375 380

Ser Met Phe Pro Ala Ala Pro Ala Pro Lys Met Ile Glu Asp Gly Lys  
385 390 395 400

Ile His Thr Val Glu His Met Ile Ser Pro Ile Ala Val Gln Pro Asp  
405 410 415

Pro Ala Ser Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala  
420 425 430

Ala Val Ile Pro Thr Val Ser Thr Pro Pro Pro Phe Gln Gly Arg Pro  
435 440 445

Ile Thr Pro Val Tyr Thr Val Ala Pro Asn Val Gln Arg Ile Pro Thr  
450 455 460

Ala Gly Ile Tyr Gly Ala Ser Tyr Val Pro Phe Ala Ala Pro Ala Thr  
465 470 475 480

Ala Thr Ile Ala Thr Leu Gln Lys Asn Ala Ala Ala Ala Ala Ala Val  
485 490 495

Tyr Gly Gly Tyr Ala Gly Tyr Ile Pro Gln Ala Phe Pro Ala Ala Ala  
500 505 510

Ile Gln Val Pro Ile Pro Asp Val Tyr Gln Thr Tyr  
515 520

&lt;210&gt; 43

&lt;211&gt; 1451

&lt;212&gt; DNA

&lt;213&gt; NM\_013952 PAX8

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (161)..(1354)

&lt;223&gt;

|   |     |
|---|-----|
| <400> 43  |     |
| ttcagaagga ggagagacac cgggccagg gcacctcgc gggcgggcgg acccaagcag   | 60  |
| tgagggcctg cagccggccg gccagggcag cggcaggcgc ggcccggacc tacgggagga | 120 |
| agccccgagc cctcggcggg ctgcgagcga ctccccggcg atg cct cac aac tcc   | 175 |
| Met Pro His Asn Ser   |     |
| 1 5   |     |
| atc aga tct ggc cat gga ggg ctg aac cag ctg gga ggg gcc ttt gtg   | 223 |
| Ile Arg Ser Gly His Gly Gly Leu Asn Gln Leu Gly Gly Ala Phe Val   |     |
| 10 15 20  |     |
| aat ggc aga cct ctg ccg gaa gtg gtc cgc cag cgc atc gta gac ctg   | 271 |
| Asn Gly Arg Pro Leu Pro Glu Val Val Arg Gln Arg Ile Val Asp Leu   |     |
| 25 30 35  |     |
| gcc cac cag ggt gta agg ccc tgc gac atc tct cgc cag ctg cgc gtc   | 319 |
| Ala His Gln Gly Val Arg Pro Cys Asp Ile Ser Arg Gln Leu Arg Val   |     |
| 40 45 50  |     |
| agc cat ggc tgc gtc agc aag atc ctt ggc agg tac tac gag act ggc   | 367 |
| Ser His Gly Cys Val Ser Lys Ile Leu Gly Arg Tyr Tyr Glu Thr Gly   |     |
| 55 60 65  |     |
| agc atc cgg cct gga gtg ata ggg ggc tcc aag ccc aag gtg gcc acc   | 415 |
| Ser Ile Arg Pro Gly Val Ile Gly Gly Ser Lys Pro Lys Val Ala Thr   |     |
| 70 75 80 85   |     |
| ccc aag gtg gtg gag aag att ggg gac tac aaa cgc cag aac cct acc   | 463 |
| Pro Lys Val Val Glu Lys Ile Gly Asp Tyr Lys Arg Gln Asn Pro Thr   |     |
| 90 95 100   |     |
| atg ttt gcc tgg gag atc cga gac cgg ctg ctg gct gag ggc gtc tgt   | 511 |
| Met Phe Ala Trp Glu Ile Arg Asp Arg Leu Leu Ala Glu Gly Val Cys   |     |
| 105 110 115   |     |
| gac aat gac act gtg ccc agt gtc agc tcc att aat aga atc atc cgg   | 559 |
| Asp Asn Asp Thr Val Pro Ser Val Ser Ser Ile Asn Arg Ile Ile Arg   |     |
| 120 125 130   |     |
| acc aaa gtg cag caa cca ttc aac ctg cct atg gac agc tgc gtg gcc   | 607 |
| Thr Lys Val Gln Gln Pro Phe Asn Leu Pro Met Asp Ser Cys Val Ala   |     |
| 135 140 145   |     |
| acc aag tcc ctg agt ccc gga cac acg ctg atc ccc agc tca gct gta   | 655 |
| Thr Lys Ser Leu Ser Pro Gly His Thr Leu Ile Pro Ser Ser Ala Val   |     |
| 150 155 160 165   |     |
| act ccc ccg gag tca ccc cag tcg gat tcc ctg ggc tcc acc tac tcc   | 703 |
| Thr Pro Pro Glu Ser Pro Gln Ser Asp Ser Leu Gly Ser Thr Tyr Ser   |     |
| 170 175 180   |     |
| atc aat ggg ctg ctg ggc atc gct cag cct ggc agc gac aag agg aaa   | 751 |
| Ile Asn Gly Leu Leu Gly Ile Ala Gln Pro Gly Ser Asp Lys Arg Lys   |     |
| 185 190 195   |     |
| atg gat gac agt gat cag gat agc tgc cga cta agc att gac tca cag   | 799 |
| Met Asp Asp Ser Asp Gln Asp Ser Cys Arg Leu Ser Ile Asp Ser Gln   |     |
| 200 205 210   |     |
| agc agc agc agc gga ccc cga aag cac ctt cgc acg gat gcc ttc agc   | 847 |
| Ser Ser Ser Ser Gly Pro Arg Lys His Leu Arg Thr Asp Ala Phe Ser   |     |
| 215 220 225   |     |

cag cac cac ctc gag ccg ctc gag tgc cca ttt gag cgg cag cac tac 895  
Gln His His Leu Glu Pro Leu Glu Cys Pro Phe Glu Arg Gln His Tyr  
230 235 240 245

cca gag gcc tat gcc tcc ccc agc cac acc aaa ggc gag cag gcc ctc 943  
Pro Glu Ala Tyr Ala Ser Pro Ser His Thr Lys Gly Glu Gln Gly Leu  
250 255 260

tac ccg ctg ccc ttg ctc aac agc acc ctg gac gac ggg aag gcc acc 991  
Tyr Pro Leu Pro Leu Leu Asn Ser Thr Leu Asp Asp Gly Lys Ala Thr  
265 270 275

ctg acc cct tcc aac acg cca ctg ggg cgc aac ctc tcg act cac cag 1039  
Leu Thr Pro Ser Asn Thr Pro Leu Gly Arg Asn Leu Ser Thr His Gln  
280 285 290

acc tac ccc gtg gtg gca gct ccg ccc ttt tgg atc tgc agc aag tcg 1087  
Thr Tyr Pro Val Val Ala Ala Pro Pro Phe Trp Ile Cys Ser Lys Ser  
295 300 305

gct ccg ggg tcc cgc cct tca atg cct ttg ccc atg ctg cct ccg tgt 1135  
Ala Pro Gly Ser Arg Pro Ser Met Pro Phe Pro Met Leu Pro Pro Cys  
310 315 320 325

acg ggc agt tca ccg gcc agg ccc tcc tct cag ggc gag aga tgg tgg 1183  
Thr Gly Ser Ser Arg Ala Arg Pro Ser Ser Gln Gly Glu Arg Trp Trp  
330 335 340

ggc cca cgc tgc ccg gat acc cac ccc aca tcc cca cca gcg gac agg 1231  
Gly Pro Arg Cys Pro Asp Thr His Pro Thr Ser Pro Pro Ala Asp Arg  
345 350 355

gca gct atg cct cct ctg cca tcg cag gca tgg tgg cag gaa gtg aat 1279  
Ala Ala Met Pro Pro Leu Pro Ser Gln Ala Trp Trp Gln Glu Val Asn  
360 365 370

act ctg gca atg cct atg gcc aca ccc cct act cct cct aca gcg agg 1327  
Thr Leu Ala Met Pro Met Ala Thr Pro Pro Thr Pro Pro Thr Ala Arg  
375 380 385

cct ggg gct tcc cca act cca gct tgc tgagttcccc atattattac 1374  
Pro Gly Ala Ser Pro Thr Pro Ala Cys  
390 395

agttccacat caaggccgag tgcaccgccc accactgcc a cggcctttga ccatctgtag 1434

ttgccatggg gacagtg 1451

<210> 44

<211> 398

<212> PRT

<213> NM\_013952 PAX8

<400> 44

Met Pro His Asn Ser Ile Arg Ser Gly His Gly Gly Leu Asn Gln Leu  
1 5 10 15

Gly Gly Ala Phe Val Asn Gly Arg Pro Leu Pro Glu Val Val Arg Gln  
20 25 30

Arg Ile Val Asp Leu Ala His Gln Gly Val Arg Pro Cys Asp Ile Ser

35

40

45

Arg Gln Leu Arg Val Ser His Gly Cys Val Ser Lys Ile Leu Gly Arg

50

55

60

Tyr Tyr Glu Thr Gly Ser Ile Arg Pro Gly Val Ile Gly Gly Ser Lys

65

70

75

80

Pro Lys Val Ala Thr Pro Lys Val Val Glu Lys Ile Gly Asp Tyr Lys

85

90

95

Arg Gln Asn Pro Thr Met Phe Ala Trp Glu Ile Arg Asp Arg Leu Leu

100

105

110

Ala Glu Gly Val Cys Asp Asn Asp Thr Val Pro Ser Val Ser Ser Ile

115

120

125

Asn Arg Ile Ile Arg Thr Lys Val Gln Gln Pro Phe Asn Leu Pro Met

130

135

140

Asp Ser Cys Val Ala Thr Lys Ser Leu Ser Pro Gly His Thr Leu Ile

145

150

155

160

Pro Ser Ser Ala Val Thr Pro Pro Glu Ser Pro Gln Ser Asp Ser Leu

165

170

175

Gly Ser Thr Tyr Ser Ile Asn Gly Leu Leu Gly Ile Ala Gln Pro Gly

180

185

190

Ser Asp Lys Arg Lys Met Asp Asp Ser Asp Gln Asp Ser Cys Arg Leu

195

200

205

Ser Ile Asp Ser Gln Ser Ser Ser Ser Gly Pro Arg Lys His Leu Arg

210

215

220

Thr Asp Ala Phe Ser Gln His His Leu Glu Pro Leu Glu Cys Pro Phe

225

230

235

240

Glu Arg Gln His Tyr Pro Glu Ala Tyr Ala Ser Pro Ser His Thr Lys

245

250

255

Gly Glu Gln Gly Leu Tyr Pro Leu Pro Leu Leu Asn Ser Thr Leu Asp

260

265

270

Asp Gly Lys Ala Thr Leu Thr Pro Ser Asn Thr Pro Leu Gly Arg Asn

275

280

285

Leu Ser Thr His Gln Thr Tyr Pro Val Val Ala Ala Pro Pro Phe Trp

290

295

300

Ile Cys Ser Lys Ser Ala Pro Gly Ser Arg Pro Ser Met Pro Phe Pro  
305 310 315 320

Met Leu Pro Pro Cys Thr Gly Ser Ser Arg Ala Arg Pro Ser Ser Gln  
325 330 335

Gly Glu Arg Trp Trp Gly Pro Arg Cys Pro Asp Thr His Pro Thr Ser  
340 345 350

Pro Pro Ala Asp Arg Ala Ala Met Pro Pro Leu Pro Ser Gln Ala Trp  
355 360 365

Trp Gln Glu Val Asn Thr Leu Ala Met Pro Met Ala Thr Pro Pro Thr  
370 375 380

Pro Pro Thr Ala Arg Pro Gly Ala Ser Pro Thr Pro Ala Cys  
385 390 395

<210> 45

<211> 326

<212> DNA

<213> AI301558 EST

<400> 45

tttctttttt gttacattca ttgattcag tccettataa accccacacc tcataaacia 60

gagattagaa actaaacaaa aaggggggag gggaaggaaa ttctagagtc gttctggttt 120

gcaggtggtt gcggtcacia agagaaatca tcaagaatgt tcaattggca tgtgtgaaag 180

attcaggggg tctgcagctg tttagtgttg atgcagtttg gtcaaaagag tatcatgtta 240

gtcttctgtg ggtttttaggg agggattatg gagcctccct cccacccac tggcttttctt 300

gtgtcacagc ctttatttct actccg 326

<210> 46

<211> 1534

<212> DNA

<213> NM\_018000 FLJ10116

<220>

<221> CDS

<222> (267)..(938)

<223>



[illegible]

His Met Gln Gly Val Leu His  
220

agaagataat tttcaaaagg gagcaatgct gtgaatgcag cttgcttctc tctacagatt 1028  
gagaagtcca gcttcaaaag ttacttgcca ctttaagcaag gaacttgta agagatcatg 1088  
gttcatgtta ctgaaaagac ttttaaggatt tgaaggta atccatagat tgctgagaac 1148  
aatggaaata tttttatitt tacagatttt gcacttctga attcaggta aaaactaact 1208  
tgtatttagt ctgcttagag gactgtgact tgaaaatttt tatataccaa tgagcttttt 1268  
ggtagcgtcc acaatgttta aaatatttca taggcgagat ccgtgttctc cttttattaa 1328  
tgcatgttag accaatttaa ctgctgtgtt tcaggaaaat tcttcttagt ttaataagca 1388  
agctaaaagt tttttttttt atatttagtg ctttaatttt gcctcatgtt atgtaaaatt 1448  
agcctgcaga ttttttctct caattctgta gactctcgca agataaacat tcaaacagtg 1508  
aaacaacaa taaaataaat aaacct 1534

<210> 47

<211> 224

<212> PRT

<213> NM\_018000 FLJ10116

<400> 47

Met Gly Leu Arg Asp Trp Leu Arg Thr Val Cys Cys Cys Cys Arg Cys  
1 5 10 15

Glu Cys Leu Glu Glu Arg Ala Leu Pro Glu Lys Glu Pro Leu Val Ser  
20 25 30

Asp Asn Asn Pro Tyr Ser Ser Phe Gly Ala Thr Leu Val Arg Asp Asp  
35 40 45

Glu Lys Asn Leu Trp Ser Met Pro His Asp Val Ser His Thr Glu Ala  
50 55 60

Asp Asp Asp Arg Thr Leu Tyr Asn Leu Ile Val Ile Arg Asn Gln Gln  
65 70 75 80

Ala Lys Asp Ser Glu Glu Trp Gln Lys Leu Asn Tyr Asp Ile His Thr  
85 90 95

Leu Arg Gln Val Arg Arg Glu Val Arg Asn Arg Trp Lys Cys Ile Leu  
100 105 110

Glu Asp Leu Gly Phe Gln Lys Glu Ala Asp Ser Leu Leu Ser Val Thr  
115 120 125

125

Lys Leu Ser Thr Ile Ser Asp Ser Lys Asn Thr Arg Lys Ala Arg Glu  
 130 135 140

Met Leu Leu Lys Leu Ala Glu Glu Thr Ser Ile Phe Pro Thr Ser Trp  
 145 150 155 160

Glu Leu Ser Glu Arg Tyr Leu Phe Val Val Asp Arg Leu Ile Ala Leu  
 165 170 175

Asp Ala Ala Glu Glu Phe Phe Lys Leu Ala Arg Arg Thr Tyr Pro Lys  
 180 185 190

Lys Pro Gly Val Pro Cys Leu Ala Asp Gly Gln Lys Glu Leu His Leu  
 195 200 205

Trp Gly Asp Leu Ser Cys Arg Leu Ala His Met Gln Gly Val Leu His  
 210 215 220

&lt;210&gt; 48

&lt;211&gt; 2385

&lt;212&gt; DNA

&lt;213&gt; NM\_144724 FLJ30532

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (71)..(1441)

&lt;223&gt;

&lt;400&gt; 48

gataccggtt tcagagtcct gggcagcgtg cgcgctcttc ctggcggctg cgcaggtgtg 60

aaaatcacia atg tca aat gat gga aga tcc agg aat cgg gac agg cgc 109  
 Met Ser Asn Asp Gly Arg Ser Arg Asn Arg Asp Arg Arg  
 1 5 10

tac gat gag gtc cca agc gac ctg ccc tat caa gat acc acc ata aga 157  
 Tyr Asp Glu Val Pro Ser Asp Leu Pro Tyr Gln Asp Thr Thr Ile Arg  
 15 20 25

acc cac cca att ctt cat gac agt gag cgg gca gtg agc gct gat ccc 205  
 Thr His Pro Ile Leu His Asp Ser Glu Arg Ala Val Ser Ala Asp Pro  
 30 35 40 45

ttg cca cca ccc cct ctc cca tta cag cca cca ttc ggc cca gac ttc 253  
 Leu Pro Pro Pro Leu Pro Leu Gln Pro Pro Phe Gly Pro Asp Phe  
 50 55 60

tac tca agt gac aca gaa gaa cca gct ata gcg cca gat ctc aaa cca 301  
 Tyr Ser Ser Asp Thr Glu Glu Pro Ala Ile Ala Pro Asp Leu Lys Pro  
 65 70 75

gta agg cgc ttt gtc cct gac tcc tgg aag aac ttt ttc aga ggg aag 349

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| Val | Arg | Arg | Phe | Val | Pro | Asp | Ser | Trp | Lys | Asn | Phe | Phe | Arg | Gly | Lys |      |
|     | 80  |     |     |     |     |     | 85  |     |     |     | 90  |     |     |     |     |      |
| aaa | aag | gac | ccc | gaa | tgg | gat | aag | ccg | gtg | tct | gat | atc | agg | tac | atc | 397  |
| Lys | Lys | Asp | Pro | Glu | Trp | Asp | Lys | Pro | Val | Ser | Asp | Ile | Arg | Tyr | Ile |      |
|     | 95  |     |     |     |     | 100 |     |     |     | 105 |     |     |     |     |     |      |
| tcc | gat | gga | gtg | gag | tgt | tca | cca | cca | gcc | tct | cca | gca | aga | cca | aac | 445  |
| Ser | Asp | Gly | Val | Glu | Cys | Ser | Pro | Pro | Ala | Ser | Pro | Ala | Arg | Pro | Asn |      |
| 110 |     |     |     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |      |
| cac | cgt | tcg | ccc | ctc | aac | tcc | tgc | aaa | gat | ccc | tac | gga | ggg | tca | gaa | 493  |
| His | Arg | Ser | Pro | Leu | Asn | Ser | Cys | Lys | Asp | Pro | Tyr | Gly | Gly | Ser | Glu |      |
|     |     |     |     | 130 |     |     |     |     | 135 |     |     |     |     | 140 |     |      |
| gga | acc | ttt | agt | tcc | cgg | aaa | gag | gct | gac | gca | gtg | ttt | ccc | cgg | gat | 541  |
| Gly | Thr | Phe | Ser | Ser | Arg | Lys | Glu | Ala | Asp | Ala | Val | Phe | Pro | Arg | Asp |      |
|     |     | 145 |     |     |     |     |     | 150 |     |     |     |     | 155 |     |     |      |
| ccc | tat | gga | tct | cta | gac | cga | cac | aca | caa | aca | gtt | cga | aca | tac | agt | 589  |
| Pro | Tyr | Gly | Ser | Leu | Asp | Arg | His | Thr | Gln | Thr | Val | Arg | Thr | Tyr | Ser |      |
|     |     | 160 |     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |      |
| gag | aag | gtg | gag | gag | tat | aac | ctg | aga | tac | tcc | tac | atg | aag | tcg | tgg | 637  |
| Glu | Lys | Val | Glu | Glu | Tyr | Asn | Leu | Arg | Tyr | Ser | Tyr | Met | Lys | Ser | Trp |      |
|     | 175 |     |     |     |     | 180 |     |     |     |     | 185 |     |     |     |     |      |
| gca | ggc | ctg | ctg | aga | ata | ctg | ggt | gtg | gtg | gag | ctg | ctt | ttg | ggg | gcc | 685  |
| Ala | Gly | Leu | Leu | Arg | Ile | Leu | Gly | Val | Val | Glu | Leu | Leu | Leu | Gly | Ala |      |
| 190 |     |     |     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |      |
| ggt | gtc | ttt | gct | tgt | gtc | aca | gct | tac | att | cac | aag | gac | agt | gag | tgg | 733  |
| Gly | Val | Phe | Ala | Cys | Val | Thr | Ala | Tyr | Ile | His | Lys | Asp | Ser | Glu | Trp |      |
|     |     |     |     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |      |
| tac | aac | ttg | ttt | gga | tat | tca | caa | ccg | tat | ggc | atg | gga | ggc | gtt | ggt | 781  |
| Tyr | Asn | Leu | Phe | Gly | Tyr | Ser | Gln | Pro | Tyr | Gly | Met | Gly | Gly | Val | Gly |      |
|     |     |     | 225 |     |     |     |     | 230 |     |     |     |     | 235 |     |     |      |
| gga | ttg | ggc | agt | atg | tat | ggg | ggc | tat | tac | tac | act | ggc | cct | aag | acc | 829  |
| Gly | Leu | Gly | Ser | Met | Tyr | Gly | Gly | Tyr | Tyr | Tyr | Thr | Gly | Pro | Lys | Thr |      |
|     |     | 240 |     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |      |
| cct | ttt | gta | ctc | gtg | gtt | gct | gga | tta | gct | tgg | atc | acc | acc | att | att | 877  |
| Pro | Phe | Val | Leu | Val | Val | Ala | Gly | Leu | Ala | Trp | Ile | Thr | Thr | Ile | Ile |      |
|     |     | 255 |     |     |     | 260 |     |     |     |     | 265 |     |     |     |     |      |
| att | ctg | gtt | ctt | ggc | atg | tcc | atg | tat | tac | cgg | acc | att | ctt | ctg | gac | 925  |
| Ile | Leu | Val | Leu | Gly | Met | Ser | Met | Tyr | Tyr | Arg | Thr | Ile | Leu | Leu | Asp |      |
|     |     |     |     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |      |
| tct | aat | tgg | tgg | ccc | cta | act | gaa | ttt | gga | att | aac | gtt | gcc | ttg | ttt | 973  |
| Ser | Asn | Trp | Trp | Pro | Leu | Thr | Glu | Phe | Gly | Ile | Asn | Val | Ala | Leu | Phe |      |
|     |     |     |     | 290 |     |     |     |     | 295 |     |     |     | 300 |     |     |      |
| att | ttg | tat | atg | gcc | gca | gcc | ata | gtc | tat | gtg | aat | gat | acc | aac | cga | 1021 |
| Ile | Leu | Tyr | Met | Ala | Ala | Ala | Ile | Val | Tyr | Val | Asn | Asp | Thr | Asn | Arg |      |
|     |     |     | 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |      |
| ggt | ggc | ctc | tgc | tac | tat | ccg | tta | ttt | aat | aca | cca | gtg | aat | gca | gtg | 1069 |
| Gly | Gly | Leu | Cys | Tyr | Tyr | Pro | Leu | Phe | Asn | Thr | Pro | Val | Asn | Ala | Val |      |
|     |     | 320 |     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |      |
| ttc | tgc | cgg | gta | gaa | gga | gga | cag | ata | gct | gca | atg | atc | ttc | ctg | ttt | 1117 |
| Phe | Cys | Arg | Val | Glu | Gly | Gly | Gln | Ile | Ala | Ala | Met | Ile | Phe | Leu | Phe |      |
|     |     | 335 |     |     |     |     | 340 |     |     |     | 345 |     |     |     |     |      |
| gtc | acc | atg | ata | gtt | tat | ctc | att | agt | gct | ttg | gtt | tgc | cta | aag | tta | 1165 |

Val Thr Met Ile Val Tyr Leu Ile Ser Ala Leu Val Cys Leu Lys Leu  
350 355 360 365

tgg agg cat gag gca gct cgg aga cat aga gaa tat atg gaa caa cag 1213  
Trp Arg His Glu Ala Ala Arg Arg His Arg Glu Tyr Met Glu Gln Gln  
370 375 380

gag ata aat gag cca tca ttg tca tcg aaa agg aaa atg tgt gaa atg 1261  
Glu Ile Asn Glu Pro Ser Leu Ser Ser Lys Arg Lys Met Cys Glu Met  
385 390 395

gcc acc agt ggt gac aga caa aga gac tca gaa gtt aat ttc aag gaa 1309  
Ala Thr Ser Gly Asp Arg Gln Arg Asp Ser Glu Val Asn Phe Lys Glu  
400 405 410

ctg aga aca gca aaa atg aaa cct gaa cta ctg agt gga cac atc ccc 1357  
Leu Arg Thr Ala Lys Met Lys Pro Glu Leu Leu Ser Gly His Ile Pro  
415 420 425

cca cgc cca gct aat ttt ttt gta ttt tta gta gag atg ggg ttt cac 1405  
Pro Arg Pro Ala Asn Phe Phe Val Phe Leu Val Glu Met Gly Phe His  
430 435 440 445

cgt gtt agc cag gat gat ctc gat ctc ctg acc tca tgatccacc 1451  
Arg Val Ser Gln Asp Asp Leu Asp Leu Leu Thr Ser  
450 455

gcctcagcct cccaaagtgt tgggattaca ggctgagtc accgcgcca gctggtattg 1511

ctttctatt cccttggac atacatgcta cagtccaca atgtagcatt tccttgaaa 1571

ctcccttttt ttttttttt gagatggagt ttctctcttg ttgccaggc tggagtacag 1631

tggatgatc ttggctcact gcagcctctg cctcctgggt tcaagcgatt ctctgcctc 1691

tgccctccaa gtagctggga ttacaggcac ccaccacat gccagctaa tttttgtat 1751

ttttagtaga gacaggattt cactatgttg gccagggttg tctcaagctc ctgacctcag 1811

atgatctacc agcctcggcc ttctgaagtg ctgggattca ggtgtgagcc actgtgcca 1871

gcagggatgc ttcatcttcc taagaattat cttggctttg gactttattc ataaatgttt 1931

tatttctgtt agtatgaaca atagactgcc ttaacaaagt ttttttttaa acaaatcgt 1991

tcttggttga ttttattcag cagcatctat catgtagata aattcccagg tgtagcatta 2051

cagcttctga ctaatatagc tgccattcag acaattaatg ttcaaaagat tttctaaagt 2111

gataaaacca aagaaaagca tgtggaaaag cagaagctta gaaagttgtg gtcactgaat 2171

gcactccctg gtttttattt gtcagtgaat tctttatgca ttcattgtta atattttaat 2231

tccatggcct tgtaggctgt gctgtgtctg aaggggtaac acctaggga acatgagcc 2291

ccttatggga ccccccaat ggaacaactt cactttctct tttatgtatt gagccctgtg 2351

ttacatttc acttaagaag agcaccagt cttt 2385

&lt;210&gt; 49

&lt;211&gt; 457

&lt;212&gt; PRT

&lt;213&gt; NM\_144724 FLJ30532

&lt;400&gt; 49

Met Ser Asn Asp Gly Arg Ser Arg Asn Arg Asp Arg Arg Tyr Asp Glu  
1 5 10 15

Val Pro Ser Asp Leu Pro Tyr Gln Asp Thr Thr Ile Arg Thr His Pro  
20 25 30

Ile Leu His Asp Ser Glu Arg Ala Val Ser Ala Asp Pro Leu Pro Pro  
35 40 45

Pro Pro Leu Pro Leu Gln Pro Pro Phe Gly Pro Asp Phe Tyr Ser Ser  
50 55 60

Asp Thr Glu Glu Pro Ala Ile Ala Pro Asp Leu Lys Pro Val Arg Arg  
65 70 75 80

Phe Val Pro Asp Ser Trp Lys Asn Phe Phe Arg Gly Lys Lys Lys Asp  
85 90 95

Pro Glu Trp Asp Lys Pro Val Ser Asp Ile Arg Tyr Ile Ser Asp Gly  
100 105 110

Val Glu Cys Ser Pro Pro Ala Ser Pro Ala Arg Pro Asn His Arg Ser  
115 120 125

Pro Leu Asn Ser Cys Lys Asp Pro Tyr Gly Gly Ser Glu Gly Thr Phe  
130 135 140

Ser Ser Arg Lys Glu Ala Asp Ala Val Phe Pro Arg Asp Pro Tyr Gly  
145 150 155 160

Ser Leu Asp Arg His Thr Gln Thr Val Arg Thr Tyr Ser Glu Lys Val  
165 170 175

Glu Glu Tyr Asn Leu Arg Tyr Ser Tyr Met Lys Ser Trp Ala Gly Leu  
180 185 190

Leu Arg Ile Leu Gly Val Val Glu Leu Leu Leu Gly Ala Gly Val Phe  
195 200 205

Ala Cys Val Thr Ala Tyr Ile His Lys Asp Ser Glu Trp Tyr Asn Leu  
210 215 220

Phe Gly Tyr Ser Gln Pro Tyr Gly Met Gly Gly Val Gly Gly Leu Gly  
225 230 235 240

Ser Met Tyr Gly Gly Tyr Tyr Tyr Thr Gly Pro Lys Thr Pro Phe Val  
245 250 255

Leu Val Val Ala Gly Leu Ala Trp Ile Thr Thr Ile Ile Ile Leu Val  
260 265 270

Leu Gly Met Ser Met Tyr Tyr Arg Thr Ile Leu Leu Asp Ser Asn Trp  
275 280 285

Trp Pro Leu Thr Glu Phe Gly Ile Asn Val Ala Leu Phe Ile Leu Tyr  
290 295 300

Met Ala Ala Ala Ile Val Tyr Val Asn Asp Thr Asn Arg Gly Gly Leu  
305 310 315 320

Cys Tyr Tyr Pro Leu Phe Asn Thr Pro Val Asn Ala Val Phe Cys Arg  
325 330 335

Val Glu Gly Gly Gln Ile Ala Ala Met Ile Phe Leu Phe Val Thr Met  
340 345 350

Ile Val Tyr Leu Ile Ser Ala Leu Val Cys Leu Lys Leu Trp Arg His  
355 360 365

Glu Ala Ala Arg Arg His Arg Glu Tyr Met Glu Gln Gln Glu Ile Asn  
370 375 380

Glu Pro Ser Leu Ser Ser Lys Arg Lys Met Cys Glu Met Ala Thr Ser  
385 390 395 400

Gly Asp Arg Gln Arg Asp Ser Glu Val Asn Phe Lys Glu Leu Arg Thr  
405 410 415

Ala Lys Met Lys Pro Glu Leu Leu Ser Gly His Ile Pro Pro Arg Pro  
420 425 430

Ala Asn Phe Phe Val Phe Leu Val Glu Met Gly Phe His Arg Val Ser  
435 440 445

Gln Asp Asp Leu Asp Leu Leu Thr Ser  
450 455

<210> 50

<211> 2280

<212> DNA

<213> NM\_006424 SLC34A2

<220>

<221> CDS

&lt;222&gt; (64)..(2133)

&lt;223&gt;

&lt;400&gt; 50

cgggccaggt ttccaggctc ggccgccgcc tccatcccag cacctgcgga gggagcgctg 60  
 acc atg gct ccc tgg cct gaa ttg gga gat gcc cag ccc aac ccc gat 108  
 Met Ala Pro Trp Pro Glu Leu Gly Asp Ala Gln Pro Asn Pro Asp  
 1 5 10 15  
 aag tac ctc gaa ggg gcc gca ggt cag cag ccc act gcc cct gat aaa 156  
 Lys Tyr Leu Glu Gly Ala Ala Gly Gln Gln Pro Thr Ala Pro Asp Lys  
 20 25 30  
 agc aaa gag acc aac aaa aca gat aac act gag gca cct gta acc aag 204  
 Ser Lys Glu Thr Asn Lys Thr Asp Asn Thr Glu Ala Pro Val Thr Lys  
 35 40 45  
 att gaa ctt ctg ccg tcc tac tcc acg gct aca ctg ata gat gag ccc 252  
 Ile Glu Leu Leu Pro Ser Tyr Ser Thr Ala Thr Leu Ile Asp Glu Pro  
 50 55 60  
 act gag gtg gat gac ccc tgg aac cta ccc act ctt cag gac tcg ggg 300  
 Thr Glu Val Asp Asp Pro Trp Asn Leu Pro Thr Leu Gln Asp Ser Gly  
 65 70 75  
 atc aag tgg tca gag aga gac acc aaa ggg aag att ctc tgt ttc ttc 348  
 Ile Lys Trp Ser Glu Arg Asp Thr Lys Gly Lys Ile Leu Cys Phe Phe  
 80 85 90 95  
 caa ggg att ggg aga ttg att tta ctt ctc gga ttt ctc tac ttt ttc 396  
 Gln Gly Ile Gly Arg Leu Ile Leu Leu Leu Gly Phe Leu Tyr Phe Phe  
 100 105 110  
 gtg tgc tcc ctg gat att ctt agt agc gcc ttc cag ctg gtt gga gga 444  
 Val Cys Ser Leu Asp Ile Leu Ser Ser Ala Phe Gln Leu Val Gly Gly  
 115 120 125  
 aaa atg gca gga cag ttc ttc agc aac agc tct att atg tcc aac cct 492  
 Lys Met Ala Gly Gln Phe Phe Ser Asn Ser Ser Ile Met Ser Asn Pro  
 130 135 140  
 ttg ttg ggg ctg gtg atc ggg gtg ctg gtg acc gtc ttg gtg cag agc 540  
 Leu Leu Gly Leu Val Ile Gly Val Leu Val Thr Val Leu Val Gln Ser  
 145 150 155  
 tcc agc acc tca acg tcc atc gtt gtc agc atg gtg tcc tct tca ttg 588  
 Ser Ser Thr Ser Thr Ser Ile Val Val Ser Met Val Ser Ser Ser Leu  
 160 165 170 175  
 ctg act gtt cgg gct gcc atc ccc att atc atg ggg gcc aac att gga 636  
 Leu Thr Val Arg Ala Ala Ile Pro Ile Ile Met Gly Ala Asn Ile Gly  
 180 185 190  
 acg tca atc acc aac act att gtt gcg ctc atg cag gtg gga gat cgg 684  
 Thr Ser Ile Thr Asn Thr Ile Val Ala Leu Met Gln Val Gly Asp Arg  
 195 200 205  
 agt gag ttc aga aga gct ttt gca gga gcc act gtc cat gac ttc ttc 732  
 Ser Glu Phe Arg Arg Ala Phe Ala Gly Ala Thr Val His Asp Phe Phe  
 210 215 220  
 aac tgg ctg tcc gtg ttg gtg ctc ttg ccc gtg gag gtg gcc acc cat 780  
 Asn Trp Leu Ser Val Leu Val Leu Leu Pro Val Glu Val Ala Thr His  
 225 230 235



|   |      |
|---|------|
| tac ctc gag atc ata acc cag ctt ata gtg gag agc ttc cac ttc aag<br>Tyr Leu Glu Ile Ile Thr Gln Leu Ile Val Glu Ser Phe His Phe Lys<br>240 245 250 255 | 828  |
| aat gga gaa gat gcc cca gat ctt ctg aaa gtc atc act aag ccc ttc<br>Asn Gly Glu Asp Ala Pro Asp Leu Leu Lys Val Ile Thr Lys Pro Phe<br>260 265 270     | 876  |
| aca aag ctc att gtc cag ctg gat aaa aaa gtt atc agc caa att gca<br>Thr Lys Leu Ile Val Gln Leu Asp Lys Lys Val Ile Ser Gln Ile Ala<br>275 280 285     | 924  |
| atg aac gat gaa aaa gcg aaa aac aag agt ctt gtc aag att tgg tgc<br>Met Asn Asp Glu Lys Ala Lys Asn Lys Ser Leu Val Lys Ile Trp Cys<br>290 295 300     | 972  |
| aaa act ttt acc aac aag acc cag att aac gtc act gtt ccc tcg act<br>Lys Thr Phe Thr Asn Lys Thr Gln Ile Asn Val Thr Val Pro Ser Thr<br>305 310 315     | 1020 |
| gct aac tgc acc tcc cct tcc ctc tgt tgg acg gat ggc atc caa aac<br>Ala Asn Cys Thr Ser Pro Ser Leu Cys Trp Thr Asp Gly Ile Gln Asn<br>320 325 330 335 | 1068 |
| tgg acc atg aag aat gtg acc tac aag gag aac atc gcc aaa tgc cag<br>Trp Thr Met Lys Asn Val Thr Tyr Lys Glu Asn Ile Ala Lys Cys Gln<br>340 345 350     | 1116 |
| cat atc ttt gtg aat ttc cac ctc ccg gat ctt gct gtg ggc acc atc<br>His Ile Phe Val Asn Phe His Leu Pro Asp Leu Ala Val Gly Thr Ile<br>355 360 365     | 1164 |
| ttg ctc ata ctc tcc ctg ctg gtc ctc tgt ggt tgc ctg atc atg att<br>Leu Leu Ile Leu Ser Leu Leu Val Leu Cys Gly Cys Leu Ile Met Ile<br>370 375 380     | 1212 |
| gtc aag atc ctg ggc tct gtg ctc aag ggg cag gtc gcc act gtc atc<br>Val Lys Ile Leu Gly Ser Val Lys Gly Gln Val Ala Thr Val Ile<br>385 390 395         | 1260 |
| aag aag acc atc aac act gat ttc ccc ttt ccc ttt gca tgg ttg act<br>Lys Lys Thr Ile Asn Thr Asp Phe Pro Phe Pro Phe Ala Trp Leu Thr<br>400 405 410 415 | 1308 |
| ggc tac ctg gcc atc ctc gtc ggg gca ggc atg acc ttc atc gta cag<br>Gly Tyr Leu Ala Ile Leu Val Gly Ala Gly Met Thr Phe Ile Val Gln<br>420 425 430     | 1356 |
| agc agc tct gtg ttc acg tcg gcc ttg acc ccc ctg att gga atc ggc<br>Ser Ser Ser Val Phe Thr Ser Ala Leu Thr Pro Leu Ile Gly Ile Gly<br>435 440 445     | 1404 |
| gtg ata acc att gag agg gct tat cca ctc acg ctg ggc tcc aac atc<br>Val Ile Thr Ile Glu Arg Ala Tyr Pro Leu Thr Leu Gly Ser Asn Ile<br>450 455 460     | 1452 |
| ggc acc acc acc acc gcc atc ctg gcc gcc tta gcc agc cct ggc aat<br>Gly Thr Thr Thr Thr Ala Ile Leu Ala Ala Leu Ala Ser Pro Gly Asn<br>465 470 475     | 1500 |
| gca ttg agg agt tca ctc cag atc gcc ctg tgc cac ttt ttc ttc aac<br>Ala Leu Arg Ser Ser Leu Gln Ile Ala Leu Cys His Phe Phe Phe Asn<br>480 485 490 495 | 1548 |
| atc tcc ggc atc ttg ctg tgg tac ccg atc ccg ttc act cgc ctg ccc<br>Ile Ser Gly Ile Leu Leu Trp Tyr Pro Ile Pro Phe Thr Arg Leu Pro<br>500 505 510     | 1596 |

atc cgc atg gcc aag ggg ctg ggc aac atc tct gcc aag tat cgc tgg 1644  
 Ile Arg Met Ala Lys Gly Leu Gly Asn Ile Ser Ala Lys Tyr Arg Trp  
 515 520 525

ttc gcc gtc ttc tac ctg atc atc ttc ttc ttc ctg atc ccg ctg acg 1692  
 Phe Ala Val Phe Tyr Leu Ile Ile Phe Phe Phe Leu Ile Pro Leu Thr  
 530 535 540

gtg ttt ggc ctc tgc ctg gcc ggc tgg cgg gtg ctg gtt ggt gtc ggg 1740  
 Val Phe Gly Leu Ser Leu Ala Gly Trp Arg Val Leu Val Gly Val Gly  
 545 550 555

gtt ccc gtc gtc ttc atc atc atc ctg gta ctg tgc ctc cga ctc ctg 1788  
 Val Pro Val Val Phe Ile Ile Ile Leu Val Leu Cys Leu Arg Leu Leu  
 560 565 570 575

cag tct cgc tgc cca cgc gtc ctg ccg aag aaa ctc cag aac tgg aac 1836  
 Gln Ser Arg Cys Pro Arg Val Leu Pro Lys Lys Leu Gln Asn Trp Asn  
 580 585 590

ttc ctg ccg ctg tgg atg cgc tgc ctg aag ccc tgg gat gcc gtc gtc 1884  
 Phe Leu Pro Leu Trp Met Arg Ser Leu Lys Pro Trp Asp Ala Val Val  
 595 600 605

tcc aag ttc acc ggc tgc ttc cag atg cgc tgc tgc tac tgc tgc cgc 1932  
 Ser Lys Phe Thr Gly Cys Phe Gln Met Arg Cys Cys Tyr Cys Cys Arg  
 610 615 620

gtg tgc tgc cgc gcg tgc tgc ttg ctg tgt ggc tgc ccc aag tgc tgc 1980  
 Val Cys Cys Arg Ala Cys Cys Leu Leu Cys Gly Cys Pro Lys Cys Cys  
 625 630 635

cgc tgc agc aag tgc tgc gag gac ttg gag gag gcg cag gag ggg cag 2028  
 Arg Cys Ser Lys Cys Cys Glu Asp Leu Glu Glu Ala Gln Glu Gly Gln  
 640 645 650 655

gat gtc cct gtc aag gct cct gag acc ttt gat aac ata acc att agc 2076  
 Asp Val Pro Val Lys Ala Pro Glu Thr Phe Asp Asn Ile Thr Ile Ser  
 660 665 670

aga gag gct cag ggt gag gtc cct gcc tgc gac tca aag acc gaa tgc 2124  
 Arg Glu Ala Gln Gly Glu Val Pro Ala Ser Asp Ser Lys Thr Glu Cys  
 675 680 685

acg gcc ttg taggggacgc cccagattgt cagggatggg gggatggtcc 2173  
 Thr Ala Leu  
 690

ttgagttttg catgctctcc tccctccac ttctgcaccc tttcaccacc tcgaggagat 2233

ttgctcccca ttagcgaatg aaattgatgc agtcctaaaa aaaaaaa 2280

&lt;210&gt; 51

&lt;211&gt; 690

&lt;212&gt; PRT

&lt;213&gt; NM\_006424 SLC34A2

&lt;400&gt; 51

Met Ala Pro Trp Pro Glu Leu Gly Asp Ala Gln Pro Asn Pro Asp Lys  
 1 5 10 15

Tyr Leu Glu Gly Ala Ala Gly Gln Gln Pro Thr Ala Pro Asp Lys Ser  
20 25 30

Lys Glu Thr Asn Lys Thr Asp Asn Thr Glu Ala Pro Val Thr Lys Ile  
35 40 45

Glu Leu Leu Pro Ser Tyr Ser Thr Ala Thr Leu Ile Asp Glu Pro Thr  
50 55 60

Glu Val Asp Asp Pro Trp Asn Leu Pro Thr Leu Gln Asp Ser Gly Ile  
65 70 75 80

Lys Trp Ser Glu Arg Asp Thr Lys Gly Lys Ile Leu Cys Phe Phe Gln  
85 90 95

Gly Ile Gly Arg Leu Ile Leu Leu Leu Gly Phe Leu Tyr Phe Phe Val  
100 105 110

Cys Ser Leu Asp Ile Leu Ser Ser Ala Phe Gln Leu Val Gly Gly Lys  
115 120 125

Met Ala Gly Gln Phe Phe Ser Asn Ser Ser Ile Met Ser Asn Pro Leu  
130 135 140

Leu Gly Leu Val Ile Gly Val Leu Val Thr Val Leu Val Gln Ser Ser  
145 150 155 160

Ser Thr Ser Thr Ser Ile Val Val Ser Met Val Ser Ser Ser Leu Leu  
165 170 175

Thr Val Arg Ala Ala Ile Pro Ile Ile Met Gly Ala Asn Ile Gly Thr  
180 185 190

Ser Ile Thr Asn Thr Ile Val Ala Leu Met Gln Val Gly Asp Arg Ser  
195 200 205

Glu Phe Arg Arg Ala Phe Ala Gly Ala Thr Val His Asp Phe Phe Asn  
210 215 220

Trp Leu Ser Val Leu Val Leu Leu Pro Val Glu Val Ala Thr His Tyr  
225 230 235 240

Leu Glu Ile Ile Thr Gln Leu Ile Val Glu Ser Phe His Phe Lys Asn  
245 250 255

Gly Glu Asp Ala Pro Asp Leu Leu Lys Val Ile Thr Lys Pro Phe Thr  
260 265 270

Lys Leu Ile Val Gln Leu Asp Lys Lys Val Ile Ser Gln Ile Ala Met

275 280 285

Asn Asp Glu Lys Ala Lys Asn Lys Ser Leu Val Lys Ile Trp Cys Lys  
290 295 300

Thr Phe Thr Asn Lys Thr Gln Ile Asn Val Thr Val Pro Ser Thr Ala  
305 310 315 320

Asn Cys Thr Ser Pro Ser Leu Cys Trp Thr Asp Gly Ile Gln Asn Trp  
325 330 335

Thr Met Lys Asn Val Thr Tyr Lys Glu Asn Ile Ala Lys Cys Gln His  
340 345 350

Ile Phe Val Asn Phe His Leu Pro Asp Leu Ala Val Gly Thr Ile Leu  
355 360 365

Leu Ile Leu Ser Leu Leu Val Leu Cys Gly Cys Leu Ile Met Ile Val  
370 375 380

Lys Ile Leu Gly Ser Val Leu Lys Gly Gln Val Ala Thr Val Ile Lys  
385 390 395 400

Lys Thr Ile Asn Thr Asp Phe Pro Phe Pro Phe Ala Trp Leu Thr Gly  
405 410 415

Tyr Leu Ala Ile Leu Val Gly Ala Gly Met Thr Phe Ile Val Gln Ser  
420 425 430

Ser Ser Val Phe Thr Ser Ala Leu Thr Pro Leu Ile Gly Ile Gly Val  
435 440 445

Ile Thr Ile Glu Arg Ala Tyr Pro Leu Thr Leu Gly Ser Asn Ile Gly  
450 455 460

Thr Thr Thr Thr Ala Ile Leu Ala Ala Leu Ala Ser Pro Gly Asn Ala  
465 470 475 480

Leu Arg Ser Ser Leu Gln Ile Ala Leu Cys His Phe Phe Phe Asn Ile  
485 490 495

Ser Gly Ile Leu Leu Trp Tyr Pro Ile Pro Phe Thr Arg Leu Pro Ile  
500 505 510

Arg Met Ala Lys Gly Leu Gly Asn Ile Ser Ala Lys Tyr Arg Trp Phe  
515 520 525

Ala Val Phe Tyr Leu Ile Ile Phe Phe Phe Leu Ile Pro Leu Thr Val  
530 535 540

Phe Gly Leu Ser Leu Ala Gly Trp Arg Val Leu Val Gly Val Gly Val

545                      550                      555                      560

Pro Val Val Phe Ile Ile Ile Leu Val Leu Cys Leu Arg Leu Leu Gln  
565                      570                      575

Ser Arg Cys Pro Arg Val Leu Pro Lys Lys Leu Gln Asn Trp Asn Phe  
580                      585                      590

Leu Pro Leu Trp Met Arg Ser Leu Lys Pro Trp Asp Ala Val Val Ser  
595                      600                      605

Lys Phe Thr Gly Cys Phe Gln Met Arg Cys Cys Tyr Cys Cys Arg Val  
610                      615                      620

Cys Cys Arg Ala Cys Cys Leu Leu Cys Gly Cys Pro Lys Cys Cys Arg  
625                      630                      635                      640

Cys Ser Lys Cys Cys Glu Asp Leu Glu Glu Ala Gln Glu Gly Gln Asp  
645                      650                      655

Val Pro Val Lys Ala Pro Glu Thr Phe Asp Asn Ile Thr Ile Ser Arg  
660                      665                      670

Glu Ala Gln Gly Glu Val Pro Ala Ser Asp Ser Lys Thr Glu Cys Thr  
675                      680                      685

Ala Leu  
690

<210> 52

<211> 529

<212> DNA

<213> AW959311, DKFZp434J037, EST

<220>

<221> misc\_feature

<222> (393)..(393)

<223> the residue at position 393 is A, T C or G

<400> 52

ctgtttcttc aatggttctc ttcccttttc catcctccaa acctggcctg agcctcctga 60

agttgctgct gtgaatctga aagacttgaa aagcctccac ctgctgtgtg gacttcactt 120

caagggggcc agcctcctct ggactccacc ttggacctca gtgactcaga actttcgcct 180

ctaagctgct ctaaagtcca gactatggat gtgttctcta ggccttcagg actctagaat 240

gtccatattt atttttatgt tcttggttt gtgttttagg aaaagtgaat ctgtctgttt 300  
 tcaataatgt gaatgctatg ttctgggaaa atccactatg acatctaagt ttgtgtaca 360  
 gagagatatt ttgcaacta ttccacett ctncacaac cccccacact ccactccaca 420  
 ctcttgagtc tctttaccta atggtcteta cctaattggac cctcgtggcc aaaaagtcca 480  
 ttaaaccaga aagggtgattg gaaaaaaaa aaaaaaaact cgagggggg 529

<210> 53

<211> 2100

<212> DNA

<213> AF111713 JAM1, junctional adhesion molecule 1

<220>

<221> CDS

<222> (287)..(1183)

<223>

<400> 53

cgaggccgag agggcgggcc agggccgcac tccggagact cgcggttgct acgcgcacca 60  
 tggctggaga cggagtttcg ctcttgatgc ccagcaggct ggagtgcgat ggcgcgattt 120  
 cggctcactg caacctcctc ctcccagagg tacttctcag ccctctagct ccaactgaga 180  
 acccagccag tcaggaagtc gctacttcgg gaacaccaac caatcagggg gccgtcacct 240  
 gctgaaggag tccttcggcg gctgttgtgt cgggagcctg atcgcg atg ggg aca 295  
 Met Gly Thr  
 1  
 aag gcg caa gtc gag agg aaa ctg ttg tgc ctc ttc ata ttg gcg atc 343  
 Lys Ala Gln Val Glu Arg Lys Leu Leu Cys Leu Phe Ile Leu Ala Ile  
 5 10 15  
 ctg ttg tgc tcc ctg gca ttg ggc agt gtt aca gtg cac tct tct gaa 391  
 Leu Leu Cys Ser Leu Ala Leu Gly Ser Val Thr Val His Ser Ser Glu  
 20 25 30 35  
 cct gaa gtc aga att cct gag aat aat cct gtg aag ttg tcc tgt gcc 439  
 Pro Glu Val Arg Ile Pro Glu Asn Asn Pro Val Lys Leu Ser Cys Ala  
 40 45 50  
 tac tcg ggc ttt tct tct ccc cgt gtg gag tgg aag ttt gac caa gga 487  
 Tyr Ser Gly Phe Ser Ser Pro Arg Val Glu Trp Lys Phe Asp Gln Gly  
 55 60 65  
 gac acc acc aga ctc gtt tgc tat aat aac aag atc aca gct tcc tat 535  
 Asp Thr Thr Arg Leu Val Cys Tyr Asn Asn Lys Ile Thr Ala Ser Tyr  
 70 75 80  
 gag gac cgg gtg acc ttc ttg cca act ggt atc acc ttc aag tcc gtg 583  
 Glu Asp Arg Val Thr Phe Leu Pro Thr Gly Ile Thr Phe Lys Ser Val  
 85 90 95

|   |      |
|---|------|
| aca cgg gaa gac act ggg aca tac act tgt atg gtc tct gag gaa ggc<br>Thr Arg Glu Asp Thr Gly Thr Tyr Thr Cys Met Val Ser Glu Glu Gly<br>100 105 110 115 | 631  |
| ggc aac agc tat ggg gag gtc aag gtc aag ctc atc gtg ctt gtg cct<br>Gly Asn Ser Tyr Gly Glu Val Lys Val Lys Leu Ile Val Leu Val Pro<br>120 125 130     | 679  |
| cca tcc aag cct aca gtt aac atc ccc tcc tct gcc acc att ggg aac<br>Pro Ser Lys Pro Thr Val Asn Ile Pro Ser Ser Ala Thr Ile Gly Asn<br>135 140 145     | 727  |
| cgg gca gtg ctg aca tgc tca gaa caa gat ggt tcc cca cct tct gaa<br>Arg Ala Val Leu Thr Cys Ser Glu Gln Asp Gly Ser Pro Pro Ser Glu<br>150 155 160     | 775  |
| tac acc tgg ttc aaa gat ggg ata gtg atg cct acg aat ccc aaa agc<br>Tyr Thr Trp Phe Lys Asp Gly Ile Val Met Pro Thr Asn Pro Lys Ser<br>165 170 175     | 823  |
| acc cgt gcc ttc agc aac tct tcc tat gtc ctg aat ccc aca aca gga<br>Thr Arg Ala Phe Ser Asn Ser Ser Tyr Val Leu Asn Pro Thr Thr Gly<br>180 185 190 195 | 871  |
| gag ctg gtc ttt gat ccc ctg tca gcc tct gat act gga gaa tac agc<br>Glu Leu Val Phe Asp Pro Leu Ser Ala Ser Asp Thr Gly Glu Tyr Ser<br>200 205 210     | 919  |
| tgt gag gca cgg aat ggg tat ggg aca ccc atg act tca aat gct gtg<br>Cys Glu Ala Arg Asn Gly Tyr Gly Thr Pro Met Thr Ser Asn Ala Val<br>215 220 225     | 967  |
| cgc atg gaa gct gtg gag cgg aat gtg ggg gtc atc gtg gca gcc gtc<br>Arg Met Glu Ala Val Glu Arg Asn Val Gly Val Ile Val Ala Ala Val<br>230 235 240     | 1015 |
| ctt gta acc ctg att ctc ctg gga atc ttg gtt ttt ggc atc tgg ttt<br>Leu Val Thr Leu Ile Leu Leu Gly Ile Leu Val Phe Gly Ile Trp Phe<br>245 250 255     | 1063 |
| gcc tat agc cga ggc cac ttt gac aga aca aag aaa ggg act tcg agt<br>Ala Tyr Ser Arg Gly His Phe Asp Arg Thr Lys Lys Gly Thr Ser Ser<br>260 265 270 275 | 1111 |
| aag aag gtg att tac agc cag cct agt gcc cga agt gaa gga gaa ttc<br>Lys Lys Val Ile Tyr Ser Gln Pro Ser Ala Arg Ser Glu Gly Glu Phe<br>280 285 290     | 1159 |
| aaa cag acc tcg tca ttc ctg gtg tgagcctggt cggtcaccg cctatcatct<br>Lys Gln Thr Ser Ser Phe Leu Val<br>295   | 1213 |
| gcatttgcct tactcaggtg ctactggact ctggcccctg atgtctgtag tttcacagga   | 1273 |
| tgccattat ttt gtcttctaca cccacaggg cccctactt cttcggtatgt gtttttaata   | 1333 |
| atgtcagcta tgtgcccatt cctccttcat gccctccctc cctttccctac cactgctgag  | 1393 |
| tggcctggaa cttgtttaaa gtgtttatc cccatttctt tgagggatca ggaaggaatc  | 1453 |
| ctgggtatgc cattgacttc ccttctaagt agacagcaaa aatggcgggg gtcgcaggaa   | 1513 |
| tctgcactca actgcccacc tggctggcag ggatccttga ataggtatct tgagcttggt   | 1573 |
| tctgggctct ttccttgtgt actgacgacc agggccagct gttctagagt gggaattaga   | 1633 |
| ggctagagcg gctgaaatgg ttgtttggtg atgacactgg ggtccttcca tctctggggc   | 1693 |

ccactctctt ctgtcttccc atgggaagtg ccactgggat ccctctgccc tgtcctcctg 1753  
 aatacaagct gactgacatt gactgtgtct gtggaaaatg ggagctcttg ttgtggagag 1813  
 catagtaaat tticagagaa cttgaagcga aaaggattta aaaccgctgc tctaaagaaa 1873  
 agaaaactgg aggetgggag cagtggctca cgctgtaat ccagaggct gaggcaggcg 1933  
 gatcacctga ggtcgggagt tcgggatcag cctgaccaac atggagaaac cctgctggaa 1993  
 atacagagtt agccaggcat ggtggtgcat gcctgtagtc ccagctgctc aggagcctgg 2053  
 caacaagagc aaaactccag ctcaaaaaaa aaaaaaaaaa aaaaaaa 2100

<210> 54

<211> 299

<212> PRT

<213> AF111713 JAM1, junctional adhesion molecule 1

<400> 54

Met Gly Thr Lys Ala Gln Val Glu Arg Lys Leu Leu Cys Leu Phe Ile  
 1 5 10 15

Leu Ala Ile Leu Leu Cys Ser Leu Ala Leu Gly Ser Val Thr Val His  
 20 25 30

Ser Ser Glu Pro Glu Val Arg Ile Pro Glu Asn Asn Pro Val Lys Leu  
 35 40 45

Ser Cys Ala Tyr Ser Gly Phe Ser Ser Pro Arg Val Glu Trp Lys Phe  
 50 55 60

Asp Gln Gly Asp Thr Thr Arg Leu Val Cys Tyr Asn Asn Lys Ile Thr  
 65 70 75 80

Ala Ser Tyr Glu Asp Arg Val Thr Phe Leu Pro Thr Gly Ile Thr Phe  
 85 90 95

Lys Ser Val Thr Arg Glu Asp Thr Gly Thr Tyr Thr Cys Met Val Ser  
 100 105 110

Glu Glu Gly Gly Asn Ser Tyr Gly Glu Val Lys Val Lys Leu Ile Val  
 115 120 125

Leu Val Pro Pro Ser Lys Pro Thr Val Asn Ile Pro Ser Ser Ala Thr  
 130 135 140

Ile Gly Asn Arg Ala Val Leu Thr Cys Ser Glu Gln Asp Gly Ser Pro  
 145 150 155 160



Pro Ser Glu Tyr Thr Trp Phe Lys Asp Gly Ile Val Met Pro Thr Asn  
165 170 175

Pro Lys Ser Thr Arg Ala Phe Ser Asn Ser Ser Tyr Val Leu Asn Pro  
180 185 190

Thr Thr Gly Glu Leu Val Phe Asp Pro Leu Ser Ala Ser Asp Thr Gly  
195 200 205

Glu Tyr Ser Cys Glu Ala Arg Asn Gly Tyr Gly Thr Pro Met Thr Ser  
210 215 220

Asn Ala Val Arg Met Glu Ala Val Glu Arg Asn Val Gly Val Ile Val  
225 230 235 240

Ala Ala Val Leu Val Thr Leu Ile Leu Leu Gly Ile Leu Val Phe Gly  
245 250 255

Ile Trp Phe Ala Tyr Ser Arg Gly His Phe Asp Arg Thr Lys Lys Gly  
260 265 270

Thr Ser Ser Lys Lys Val Ile Tyr Ser Gln Pro Ser Ala Arg Ser Glu  
275 280 285

Gly Glu Phe Lys Gln Thr Ser Ser Phe Leu Val  
290 295

<210> 55

<211> 2154

<212> DNA

<213> NM\_006636 MTHFD2, methylene tetrahydrofolate

<220>

<221> CDS

<222> (77)..(1108)

<223>

<400> 55

atataaccgc gtggcccgcg cgcgcgcttc cctcccgcg cagtcaccgg cgcggtctat 60

ggctgcgact tctcta atg tct gct ttg gct gcc cgg ctg ctg cag ccc gcg 112  
Met Ser Ala Leu Ala Ala Arg Leu Leu Gln Pro Ala  
1 5 10

cac agc tgc tcc ctt cgc ctt cgc cct ttc cac ctc gcg gca gtt cga 160  
His Ser Cys Ser Leu Arg Leu Arg Pro Phe His Leu Ala Ala Val Arg  
15 20 25

|   |     |
|---|-----|
| aat gaa gct gtt gtc att tct gga agg aaa ctg gcc cag cag atc aag<br>Asn Glu Ala Val Val Ile Ser Gly Arg Lys Leu Ala Gln Gln Ile Lys<br>30 35 40        | 208 |
| cag gaa gtg cgg cag gag gta gaa gag tgg gtg gcc tca ggc aac aaa<br>Gln Glu Val Arg Gln Glu Val Glu Glu Trp Val Ala Ser Gly Asn Lys<br>45 50 55 60     | 256 |
| cgg cca cac ctg agt gtg atc ctg gtt ggc gag aat cct gca agt cac<br>Arg Pro His Leu Ser Val Ile Leu Val Gly Glu Asn Pro Ala Ser His<br>65 70 75        | 304 |
| tcc tat gtc ctc aac aaa acc agg gca gct gca gtt gtg gga atc aac<br>Ser Tyr Val Leu Asn Lys Thr Arg Ala Ala Val Val Gly Ile Asn<br>80 85 90            | 352 |
| agt gag aca att atg aaa cca gct tca att tca gag gaa gaa ttg ttg<br>Ser Glu Thr Ile Met Lys Pro Ala Ser Ile Ser Glu Glu Glu Leu Leu<br>95 100 105      | 400 |
| aat tta atc aat aaa ctg aat aat gat gat aat gta gat ggc ctc ctt<br>Asn Leu Ile Asn Lys Leu Asn Asn Asp Asp Asn Val Asp Gly Leu Leu<br>110 115 120     | 448 |
| gtt cag ttg cct ctt cca gag cat att gat gag aga agg atc tgc aat<br>Val Gln Leu Pro Leu Pro Glu His Ile Asp Glu Arg Arg Ile Cys Asn<br>125 130 135 140 | 496 |
| gct gtt tct cca gac aag gat gtt gat ggc ttt cat gta att aat gta<br>Ala Val Ser Pro Asp Lys Asp Val Asp Gly Phe His Val Ile Asn Val<br>145 150 155     | 544 |
| gga cga atg tgt ttg gat cag tat tcc atg tta ccg gct act cca tgg<br>Gly Arg Met Cys Leu Asp Gln Tyr Ser Met Leu Pro Ala Thr Pro Trp<br>160 165 170     | 592 |
| ggg gtg tgg gaa ata atc aag cga act ggc att cca acc cta ggg aag<br>Gly Val Trp Glu Ile Ile Lys Arg Thr Gly Ile Pro Thr Leu Gly Lys<br>175 180 185     | 640 |
| aat gtg gtt gtg gct gga agg tca aaa aac gtt gga atg ccc att gca<br>Asn Val Val Val Ala Gly Arg Ser Lys Asn Val Gly Met Pro Ile Ala<br>190 195 200     | 688 |
| atg tta ctg cac aca gat ggg gcg cat gaa cgt ccc gga ggt gat gcc<br>Met Leu Leu His Thr Asp Gly Ala His Glu Arg Pro Gly Gly Asp Ala<br>205 210 215 220 | 736 |
| act gtt aca ata tct cat cga tat act ccc aaa gag cag ttg aag aaa<br>Thr Val Thr Ile Ser His Arg Tyr Thr Pro Lys Glu Gln Leu Lys Lys<br>225 230 235     | 784 |
| cat aca att ctt gca gat att gta ata tct gct gca ggt att cca aat<br>His Thr Ile Leu Ala Asp Ile Val Ile Ser Ala Ala Gly Ile Pro Asn<br>240 245 250     | 832 |
| ctg atc aca gca gat atg atc aag gaa gga gca gca gtc att gat gtg<br>Leu Ile Thr Ala Asp Met Ile Lys Glu Gly Ala Ala Val Ile Asp Val<br>255 260 265     | 880 |
| gga ata aat aga gtt cac gat cct gta act gcc aaa ccc aag ttg gtt<br>Gly Ile Asn Arg Val His Asp Pro Val Thr Ala Lys Pro Lys Leu Val<br>270 275 280     | 928 |
| gga gat gtg gat ttt gaa gga gtc aga caa aaa gct ggg tat atc act<br>Gly Asp Val Asp Phe Glu Gly Val Arg Gln Lys Ala Gly Tyr Ile Thr<br>285 290 295 300 | 976 |

cca gtt cct gga ggt gtt ggc ccc atg aca gtg gca atg cta atg aag 1024  
 Pro Val Pro Gly Gly Val Gly Pro Met Thr Val Ala Met Leu Met Lys  
                   305                  310                  315

aat acc att att gct gca aaa aag gtg ctg agg ctt gaa gag cga gaa 1072  
 Asn Thr Ile Ile Ala Ala Lys Lys Val Leu Arg Leu Glu Glu Arg Glu  
                   320                  325                  330

gtg ctg aag tct aaa gag ctt ggg gta gcc act aat taactactgt 1118  
 Val Leu Lys Ser Lys Glu Leu Gly Val Ala Thr Asn  
                   335                  340

gtcttctgtg tcacaaacag cactccaggc cagctcaaga agcaaagcag gccaatagaa 1178

atgcaatatt ttttaatttat tctactgaaa tgggttaaaa tgatgccttg tatttattga 1238

aagcttaaat ggggtgggtgt ttctgcacat acctctgcag tacctcacca gggagcattc 1298

cagtatcatg cagggtcctg tgatctagcc aggagcagcc attaacctag tgattaatat 1358

gggagacatt accatatgga ggatggatgc ttcactttgt caagcacctc agttacacat 1418

tcgccttttc taggattgca ttcccaagt gctattgcaa taacagtga tactcatttt 1478

aggtagcaga ccttttgagt tcaactgac aaaccaaagg aaaagtgtg ctagagaaaa 1538

ttggggaaaa ggtgaaaaag aaaaaatggt agtaattgag cagaaaaaaa ttaatttata 1598

tatgtattga ttggcaacca gatttatcta agtagaactg aattggctag gaaaaagaa 1658

aaactgcatg ttaatcattt tcctaagctg tccttttgag gcttagtcag tttattggga 1718

aaatgtttag gattatctct tgetattagt actcatttta tgtatgttac ccttcagtaa 1778

gttctcccca ttttagtttt ctaggactga aaggattctt ttctacatta tacatgtgtg 1838

ttgtcatatt tggtttttgc tatatacttt aacttcattg ttaaattttt gtattgtata 1898

gtttcttttg tgtatcttaa aacctatttt tgaaaaacaa acttggttg ataactattt 1958

gggcagcttg ggtaagtacg caacttactt ttccaccaa gaactgtcag cagctgcctg 2018

cttttctgtg atgtatgtat cctgttgact ttccagaaa ttttttaaga gtttgagtta 2078

ctattgaatt taatcagact ttctgattaa agggttttct ttctttttta ataaaacaca 2138

tctgtctggt atggta 2154

&lt;210&gt; 56

&lt;211&gt; 344

&lt;212&gt; PRT

&lt;213&gt; NM\_006636 MTHFD2, methylene tetrahydrofolate

&lt;400&gt; 56

Met Ser Ala Leu Ala Ala Arg Leu Leu Gln Pro Ala His Ser Cys Ser  
 1                  5                  10                  15

Leu Arg Leu Arg Pro Phe His Leu Ala Ala Val Arg Asn Glu Ala Val  
 20                  25                  30

Val Ile Ser Gly Arg Lys Leu Ala Gln Gln Ile Lys Gln Glu Val Arg  
35 40 45

Gln Glu Val Glu Glu Trp Val Ala Ser Gly Asn Lys Arg Pro His Leu  
50 55 60

Ser Val Ile Leu Val Gly Glu Asn Pro Ala Ser His Ser Tyr Val Leu  
65 70 75 80

Asn Lys Thr Arg Ala Ala Ala Val Val Gly Ile Asn Ser Glu Thr Ile  
85 90 95

Met Lys Pro Ala Ser Ile Ser Glu Glu Glu Leu Leu Asn Leu Ile Asn  
100 105 110

Lys Leu Asn Asn Asp Asp Asn Val Asp Gly Leu Leu Val Gln Leu Pro  
115 120 125

Leu Pro Glu His Ile Asp Glu Arg Arg Ile Cys Asn Ala Val Ser Pro  
130 135 140

Asp Lys Asp Val Asp Gly Phe His Val Ile Asn Val Gly Arg Met Cys  
145 150 155 160

Leu Asp Gln Tyr Ser Met Leu Pro Ala Thr Pro Trp Gly Val Trp Glu  
165 170 175

Ile Ile Lys Arg Thr Gly Ile Pro Thr Leu Gly Lys Asn Val Val Val  
180 185 190

Ala Gly Arg Ser Lys Asn Val Gly Met Pro Ile Ala Met Leu Leu His  
195 200 205

Thr Asp Gly Ala His Glu Arg Pro Gly Gly Asp Ala Thr Val Thr Ile  
210 215 220

Ser His Arg Tyr Thr Pro Lys Glu Gln Leu Lys Lys His Thr Ile Leu  
225 230 235 240

Ala Asp Ile Val Ile Ser Ala Ala Gly Ile Pro Asn Leu Ile Thr Ala  
245 250 255

Asp Met Ile Lys Glu Gly Ala Ala Val Ile Asp Val Gly Ile Asn Arg  
260 265 270

Val His Asp Pro Val Thr Ala Lys Pro Lys Leu Val Gly Asp Val Asp  
275 280 285

Phe Glu Gly Val Arg Gln Lys Ala Gly Tyr Ile Thr Pro Val Pro Gly  
290 295 300

Gly Val Gly Pro Met Thr Val Ala Met Leu Met Lys Asn Thr Ile Ile  
305 310 315 320

Ala Ala Lys Lys Val Leu Arg Leu Glu Glu Arg Glu Val Leu Lys Ser  
325 330 335

Lys Glu Leu Gly Val Ala Thr Asn  
340

<210> 57

<211> 1117

<212> DNA

<213> NM\_006149 galectin 4, LGALS4

<220>

<221> CDS

<222> (57)..(1025)

<223>

<400> 57

atctcccact cctgcagctc ttctcacagg accagccact agcgcagcct cgagcgc atg 59  
Met  
1

gcc tat gtc ccc gca ccg ggc tac cag ccc acc tac aac ccg acg ctg 107  
Ala Tyr Val Pro Ala Pro Gly Tyr Gln Pro Thr Tyr Asn Pro Thr Leu  
5 10 15

cct tac tac cag ccc atc ccg ggc ggg ctc aac gtg gga atg tct gtt 155  
Pro Tyr Tyr Gln Pro Ile Pro Gly Gly Leu Asn Val Gly Met Ser Val  
20 25 30

tac atc caa gga gtg gcc agc gag cac atg aag cgg ttc ttc gtg aac 203  
Tyr Ile Gln Gly Val Ala Ser Glu His Met Lys Arg Phe Phe Val Asn  
35 40 45

ttt gtg gtt ggg cag gat ccg ggc tca gac gtc gcc ttc cac ttc aat 251  
Phe Val Val Gly Gln Asp Pro Gly Ser Asp Val Ala Phe His Phe Asn  
50 55 60 65

ccg cgg ttt gac ggc tgg gac aag gtg gtc ttc aac acg ttg cag ggc 299  
Pro Arg Phe Asp Gly Trp Asp Lys Val Val Phe Asn Thr Leu Gln Gly  
70 75 80

ggg aag tgg ggc agc gag gag agg aag agg agc atg ccc ttc aaa aag 347  
Gly Lys Trp Gly Ser Glu Glu Arg Lys Arg Ser Met Pro Phe Lys Lys  
85 90 95

ggg gcc gcc ttt gag ctg gtc ttc ata gtc ctg gct gag cac tac aag 395  
Gly Ala Ala Phe Glu Leu Val Phe Ile Val Leu Ala Glu His Tyr Lys  
100 105 110

gtg gtg gta aat gga aat ccc ttc tat gag tac ggg cac cgg ctt ccc 443  
Val Val Val Asn Gly Asn Pro Phe Tyr Glu Tyr Gly His Arg Leu Pro

| 115   | 120 | 125 |      |
|---|-----|-----|------|
| cta cag atg gtc acc cac ctg caa gtg gat ggg gat ctg caa ctt caa |     |     | 491  |
| Leu Gln Met Val Thr His Leu Gln Val Asp Gly Asp Leu Gln Leu Gln |     |     |      |
| 130   | 135 | 140 | 145  |
| tca atc aac ttc atc gga ggc cag ccc ctc cgg ccc cag gga ccc ccg |     |     | 539  |
| Ser Ile Asn Phe Ile Gly Gly Gln Pro Leu Arg Pro Gln Gly Pro Pro |     |     |      |
| 150   | 155 | 160 |      |
| atg atg cca cct tac cct ggt ccc gga cat tgc cat caa cag ctg aac |     |     | 587  |
| Met Met Pro Pro Tyr Pro Gly Pro Gly His Cys His Gln Gln Leu Asn |     |     |      |
| 165   | 170 | 175 |      |
| agc ctg ccc acc atg gaa gga ccc cca acc ttc aac ccg cct gtg cca |     |     | 635  |
| Ser Leu Pro Thr Met Glu Gly Pro Pro Thr Phe Asn Pro Pro Val Pro |     |     |      |
| 180   | 185 | 190 |      |
| tat ttc ggg agg ctg caa gga ggg ctc aca gct cga aga acc atc atc |     |     | 683  |
| Tyr Phe Gly Arg Leu Gln Gly Gly Leu Thr Ala Arg Arg Thr Ile Ile |     |     |      |
| 195   | 200 | 205 |      |
| atc aag ggc tat gtg cct ccc aca ggc aag agc ttt gct atc aac ttc |     |     | 731  |
| Ile Lys Gly Tyr Val Pro Pro Thr Gly Lys Ser Phe Ala Ile Asn Phe |     |     |      |
| 210   | 215 | 220 | 225  |
| aag gtg ggc tcc tca ggg gac ata gct ctg cac att aat ccc cgc atg |     |     | 779  |
| Lys Val Gly Ser Ser Gly Asp Ile Ala Leu His Ile Asn Pro Arg Met |     |     |      |
| 230   | 235 | 240 |      |
| ggc aac ggt acc gtg gtc cgg aac agc ctt ctg aat ggc tcg tgg gga |     |     | 827  |
| Gly Asn Gly Thr Val Val Arg Asn Ser Leu Leu Asn Gly Ser Trp Gly |     |     |      |
| 245   | 250 | 255 |      |
| tcc gag gag aag aag atc acc cac aac cca ttt ggt ccc gga cag ttc |     |     | 875  |
| Ser Glu Glu Lys Lys Ile Thr His Asn Pro Phe Gly Pro Gly Gln Phe |     |     |      |
| 260   | 265 | 270 |      |
| ttt gat ctg tcc att cgc tgt ggc ttg gat cgc ttc aag gtt tac gcc |     |     | 923  |
| Phe Asp Leu Ser Ile Arg Cys Gly Leu Asp Arg Phe Lys Val Tyr Ala |     |     |      |
| 275   | 280 | 285 |      |
| aat ggc cag cac ctc ttt gac ttt gcc cat cgc ctc tcg gcc ttc cag |     |     | 971  |
| Asn Gly Gln His Leu Phe Asp Phe Ala His Arg Leu Ser Ala Phe Gln |     |     |      |
| 290   | 295 | 300 | 305  |
| agg gtg gac aca ttg gaa atc cag ggt gat gtc acc ttg tcc tat gtc |     |     | 1019 |
| Arg Val Asp Thr Leu Glu Ile Gln Gly Asp Val Thr Leu Ser Tyr Val |     |     |      |
| 310   | 315 | 320 |      |
| cag atc taatctattc ctggggccat aatcatggg aaaacagaat tatccctag    |     |     | 1075 |
| Gln Ile   |     |     |      |
| gactcctttc taagccccta ataaaatgtc tgagggtgtc tc                  |     |     | 1117 |

&lt;210&gt; 58

&lt;211&gt; 323

&lt;212&gt; PRT

&lt;213&gt; NM\_006149 galectin 4, LGALS4

&lt;400&gt; 58

Met Ala Tyr Val Pro Ala Pro Gly Tyr Gln Pro Thr Tyr Asn Pro Thr  
1 5 10 15

Leu Pro Tyr Tyr Gln Pro Ile Pro Gly Gly Leu Asn Val Gly Met Ser  
20 25 30

Val Tyr Ile Gln Gly Val Ala Ser Glu His Met Lys Arg Phe Phe Val  
35 40 45

Asn Phe Val Val Gly Gln Asp Pro Gly Ser Asp Val Ala Phe His Phe  
50 55 60

Asn Pro Arg Phe Asp Gly Trp Asp Lys Val Val Phe Asn Thr Leu Gln  
65 70 75 80

Gly Gly Lys Trp Gly Ser Glu Glu Arg Lys Arg Ser Met Pro Phe Lys  
85 90 95

Lys Gly Ala Ala Phe Glu Leu Val Phe Ile Val Leu Ala Glu His Tyr  
100 105 110

Lys Val Val Val Asn Gly Asn Pro Phe Tyr Glu Tyr Gly His Arg Leu  
115 120 125

Pro Leu Gln Met Val Thr His Leu Gln Val Asp Gly Asp Leu Gln Leu  
130 135 140

Gln Ser Ile Asn Phe Ile Gly Gly Gln Pro Leu Arg Pro Gln Gly Pro  
145 150 155 160

Pro Met Met Pro Pro Tyr Pro Gly Pro Gly His Cys His Gln Gln Leu  
165 170 175

Asn Ser Leu Pro Thr Met Glu Gly Pro Pro Thr Phe Asn Pro Pro Val  
180 185 190

Pro Tyr Phe Gly Arg Leu Gln Gly Gly Leu Thr Ala Arg Arg Thr Ile  
195 200 205

Ile Ile Lys Gly Tyr Val Pro Pro Thr Gly Lys Ser Phe Ala Ile Asn  
210 215 220

Phe Lys Val Gly Ser Ser Gly Asp Ile Ala Leu His Ile Asn Pro Arg  
225 230 235 240

Met Gly Asn Gly Thr Val Val Arg Asn Ser Leu Leu Asn Gly Ser Trp  
245 250 255

Gly Ser Glu Glu Lys Lys Ile Thr His Asn Pro Phe Gly Pro Gly Gln  
260 265 270

Phe Phe Asp Leu Ser Ile Arg Cys Gly Leu Asp Arg Phe Lys Val Tyr  
275 280 285

Ala Asn Gly Gln His Leu Phe Asp Phe Ala His Arg Leu Ser Ala Phe  
290 295 300

Gln Arg Val Asp Thr Leu Glu Ile Gln Gly Asp Val Thr Leu Ser Tyr  
305 310 315 320

Val Gln Ile

<210> 59

<211> 3697

<212> DNA

<213> NM\_004063 cadherin 17, CDH17

<220>

<221> CDS

<222> (121)..(2616)

<223>

<400> 59

aggagtggtt cccgggggag atactccagt cgtagcaaga gtctcgacca ctgaatggaa 60

gaaaaggact ttttaaccacc attttgtgac ttacagaaag gaatttgaat aaagaaaact 120

atg ata ctt cag gcc cat ctt cac tcc ctg tgt ctt ctt atg ctt tat 168  
Met Ile Leu Gln Ala His Leu His Ser Leu Cys Leu Leu Met Leu Tyr  
1 5 10 15

ttg gca act gga tat ggc caa gag ggg aag ttt agt gga ccc ctg aaa 216  
Leu Ala Thr Gly Tyr Gly Gln Glu Gly Lys Phe Ser Gly Pro Leu Lys  
20 25 30

ccc atg aca ttt tct att tat gaa ggc caa gaa ccg agt caa att ata 264  
Pro Met Thr Phe Ser Ile Tyr Glu Gly Gln Glu Pro Ser Gln Ile Ile  
35 40 45

ttc cag ttt aag gcc aat cct cct gct gtg act ttt gaa cta act ggg 312  
Phe Gln Phe Lys Ala Asn Pro Pro Ala Val Thr Phe Glu Leu Thr Gly  
50 55 60

gag aca gac aac ata ttt gtg ata gaa cgg gag gga ctt ctg tat tac 360  
Glu Thr Asp Asn Ile Phe Val Ile Glu Arg Glu Gly Leu Leu Tyr Tyr  
65 70 75 80

aac aga gcc ttg gac agg gaa aca aga tct act cac aat ctc cag gtt 408  
Asn Arg Ala Leu Asp Arg Glu Thr Arg Ser Thr His Asn Leu Gln Val  
85 90 95

gca gcc ctg gac gct aat gga att ata gtg gag ggt cca gtc cct atc 456  
Ala Ala Leu Asp Ala Asn Gly Ile Ile Val Glu Gly Pro Val Pro Ile



| 100   | 105 | 110 |      |
|---|-----|-----|------|
| acc ata gaa gtg aag gac atc aac gac aat cga ccc acg ttt ctc cag<br>Thr Ile Glu Val Lys Asp Ile Asn Asp Asn Arg Pro Thr Phe Leu Gln<br>115 120 125     |     |     | 504  |
| tca aag tac gaa ggc tca gta agg cag aac tct cgc cca gga aag ccc<br>Ser Lys Tyr Glu Gly Ser Val Arg Gln Asn Ser Arg Pro Gly Lys Pro<br>130 135 140     |     |     | 552  |
| ttc ttg tat gtc aat gcc aca gac ctg gat gat cgc gcc act ccc aat<br>Phe Leu Tyr Val Asn Ala Thr Asp Leu Asp Asp Pro Ala Thr Pro Asn<br>145 150 155 160 |     |     | 600  |
| ggc cag ctt tat tac cag att gtc atc cag ctt ccc atg atc aac aat<br>Gly Gln Leu Tyr Tyr Gln Ile Val Ile Gln Leu Pro Met Ile Asn Asn<br>165 170 175     |     |     | 648  |
| gtc atg tac ttt cag atc aac aac aaa acg gga gcc atc tct ctt acc<br>Val Met Tyr Phe Gln Ile Asn Asn Lys Thr Gly Ala Ile Ser Leu Thr<br>180 185 190     |     |     | 696  |
| cga gag gga tct cag gaa ttg aat cct gct aag aat cct tcc tat aat<br>Arg Glu Gly Ser Gln Glu Leu Asn Pro Ala Lys Asn Pro Ser Tyr Asn<br>195 200 205     |     |     | 744  |
| ctg gtg atc tca gtg aag gac atg gga ggc cag agt gag aat tcc ttc<br>Leu Val Ile Ser Val Lys Asp Met Gly Gly Gln Ser Glu Asn Ser Phe<br>210 215 220     |     |     | 792  |
| agt gat acc aca tct gtg gat atc ata gtg aca gag aat att tgg aaa<br>Ser Asp Thr Thr Ser Val Asp Ile Ile Val Thr Glu Asn Ile Trp Lys<br>225 230 235 240 |     |     | 840  |
| gca cca aaa cct gtg gag atg gtg gaa aac tca act gat cct cac ccc<br>Ala Pro Lys Pro Val Glu Met Val Glu Asn Ser Thr Asp Pro His Pro<br>245 250 255     |     |     | 888  |
| atc aaa atc act cag gtg cgg tgg aat gat ccc ggt gca caa tat tcc<br>Ile Lys Ile Thr Gln Val Arg Trp Asn Asp Pro Gly Ala Gln Tyr Ser<br>260 265 270     |     |     | 936  |
| tta gtt gac aaa gag aag ctg cca aga ttc cca ttt tca att gac cag<br>Leu Val Asp Lys Glu Lys Leu Pro Arg Phe Pro Phe Ser Ile Asp Gln<br>275 280 285     |     |     | 984  |
| gaa gga gat att tac gtg act cag ccc ttg gac cga gaa gaa aag gat<br>Glu Gly Asp Ile Tyr Val Thr Gln Pro Leu Asp Arg Glu Glu Lys Asp<br>290 295 300     |     |     | 1032 |
| gca tat gtt ttt tat gca gtt gca aag gat gag tac gga aaa cca ctt<br>Ala Tyr Val Phe Tyr Ala Val Ala Lys Asp Glu Tyr Gly Lys Pro Leu<br>305 310 315 320 |     |     | 1080 |
| tca tat ccg ctg gaa att cat gta aaa gtt aaa gat att aat gat aat<br>Ser Tyr Pro Leu Glu Ile His Val Lys Val Lys Asp Ile Asn Asp Asn<br>325 330 335     |     |     | 1128 |
| cca cct aca tgt ccg tca cca gta acc gta ttt gag gtc cag gag aat<br>Pro Pro Thr Cys Pro Ser Pro Val Thr Val Phe Glu Val Gln Glu Asn<br>340 345 350     |     |     | 1176 |
| gaa cga ctg ggt aac agt atc ggg acc ctt act gca cat gac agg gat<br>Glu Arg Leu Gly Asn Ser Ile Gly Thr Leu Thr Ala His Asp Arg Asp<br>355 360 365     |     |     | 1224 |
| gaa gaa aat act gcc aac agt ttt cta aac tac agg att gtg gag caa<br>Glu Glu Asn Thr Ala Asn Ser Phe Leu Asn Tyr Arg Ile Val Glu Gln                    |     |     | 1272 |

| 370   | 375 | 380 |      |
|---|-----|-----|------|
| act ccc aaa ctt ccc atg gat gga ctc ttc cta atc caa acc tat gct<br>Thr Pro Lys Leu Pro Met Asp Gly Leu Phe Leu Ile Gln Thr Tyr Ala<br>385 390 395 400 |     |     | 1320 |
| gga atg tta cag tta gct aaa cag tcc ttg aag aag caa gat act cct<br>Gly Met Leu Gln Leu Ala Lys Gln Ser Leu Lys Lys Gln Asp Thr Pro<br>405 410 415     |     |     | 1368 |
| cag tac aac tta acg ata gag gtg tct gac aaa gat ttc aag acc ctt<br>Gln Tyr Asn Leu Thr Ile Glu Val Ser Asp Lys Asp Phe Lys Thr Leu<br>420 425 430     |     |     | 1416 |
| tgt ttt gtg caa atc aac gtt att gat atc aat gat cag atc ccc atc<br>Cys Phe Val Gln Ile Asn Val Ile Asp Ile Asn Asp Gln Ile Pro Ile<br>435 440 445     |     |     | 1464 |
| ttt gaa aaa tca gat tat gga aac ctg act ctt gct gaa gac aca aac<br>Phe Glu Lys Ser Asp Tyr Gly Asn Leu Thr Leu Ala Glu Asp Thr Asn<br>450 455 460     |     |     | 1512 |
| att ggg tcc acc atc tta acc atc cag gcc act gat gct gat gag cca<br>Ile Gly Ser Thr Ile Leu Thr Ile Gln Ala Thr Asp Ala Asp Glu Pro<br>465 470 475 480 |     |     | 1560 |
| ttt act ggg agt tct aaa att ctg tat cat atc ata aag gga gac agt<br>Phe Thr Gly Ser Ser Lys Ile Leu Tyr His Ile Ile Lys Gly Asp Ser<br>485 490 495     |     |     | 1608 |
| gag gga cgc ctg ggg gtt gac aca gat ccc cat acc aac acc gga tat<br>Glu Gly Arg Leu Gly Val Asp Thr Asp Pro His Thr Asn Thr Gly Tyr<br>500 505 510     |     |     | 1656 |
| gtc ata att aaa aag cct ctt gat ttt gaa aca gca gct gtt tcc aac<br>Val Ile Ile Lys Lys Pro Leu Asp Phe Glu Thr Ala Ala Val Ser Asn<br>515 520 525     |     |     | 1704 |
| att gtg ttc aaa gca gaa aat cct gag cct cta gtg ttt ggt gtg aag<br>Ile Val Phe Lys Ala Glu Asn Pro Glu Pro Leu Val Phe Gly Val Lys<br>530 535 540     |     |     | 1752 |
| tac aat gca agt tct ttt gcc aag ttc acg ctt att gtg aca gat gtg<br>Tyr Asn Ala Ser Ser Phe Ala Lys Phe Thr Leu Ile Val Thr Asp Val<br>545 550 555 560 |     |     | 1800 |
| aat gaa gca cct caa ttt tcc caa cac gta ttc caa gcg aaa gtc agt<br>Asn Glu Ala Pro Gln Phe Ser Gln His Val Phe Gln Ala Lys Val Ser<br>565 570 575     |     |     | 1848 |
| gag gat gta gct ata ggc act aaa gtg ggc aat gtg act gcc aag gat<br>Glu Asp Val Ala Ile Gly Thr Lys Val Gly Asn Val Thr Ala Lys Asp<br>580 585 590     |     |     | 1896 |
| cca gaa ggt ctg gac ata agc tat tca ctg agg gga gac aca aga ggt<br>Pro Glu Gly Leu Asp Ile Ser Tyr Ser Leu Arg Gly Asp Thr Arg Gly<br>595 600 605     |     |     | 1944 |
| tgg ctt aaa att gac cac gtg act ggt gag atc ttt agt gtg gct cca<br>Trp Leu Lys Ile Asp His Val Thr Gly Glu Ile Phe Ser Val Ala Pro<br>610 615 620     |     |     | 1992 |
| ttg gac aga gaa gcc gga agt cca tat cgg gta caa gtg gtg gcc aca<br>Leu Asp Arg Glu Ala Gly Ser Pro Tyr Arg Val Gln Val Val Ala Thr<br>625 630 635 640 |     |     | 2040 |
| gaa gta ggg ggg tct tcc ttg agc tct gtg tca gag ttc cac ctg atc<br>Glu Val Gly Gly Ser Ser Leu Ser Ser Val Ser Glu Phe His Leu Ile<br>645 650 655 660 |     |     | 2088 |

|                                   |                                   |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
|-----------------------------------|-----------------------------------|------|--|--|--|--|--|--|--|-----|--|--|--|--|--|--|--|--|--|-----|--|--|--|--|--|--|--|--|--|--|
| 645                               |                                   |      |  |  |  |  |  |  |  | 650 |  |  |  |  |  |  |  |  |  | 655 |  |  |  |  |  |  |  |  |  |  |
| ctt atg gat gtg aat gac aac cct   | ccc agg cta gcc aag gac tac acg   | 2136 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| Leu Met Asp Val Asn Asp Asn Pro   | Pro Arg Leu Ala Lys Asp Tyr Thr   |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| 660                               | 665 670                           |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| ggc ttg ttc ttc tgc cat ccc ctc   | agt gca cct gga agt ctc att ttc   | 2184 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| Gly Leu Phe Phe Cys His Pro Leu   | Ser Ala Pro Gly Ser Leu Ile Phe   |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| 675                               | 680 685                           |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| gag gct act gat gat gat cag cac   | tta ttt cgg ggt ccc cat ttt aca   | 2232 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| Glu Ala Thr Asp Asp Asp Gln His   | Leu Phe Arg Gly Pro His Phe Thr   |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| 690                               | 695 700                           |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| ttt tcc ctc ggc agt gga agc tta   | caa aac gac tgg gaa gtt tcc aaa   | 2280 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| Phe Ser Leu Gly Ser Gly Ser Leu   | Gln Asn Asp Trp Glu Val Ser Lys   |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| 705                               | 710 715 720                       |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| atc aat ggt act cat gcc cga ctg   | tct acc agg cac aca gag ttt gag   | 2328 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| Ile Asn Gly Thr His Ala Arg Leu   | Ser Thr Arg His Thr Glu Phe Glu   |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| 725                               | 730 735                           |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| gag agg gag tat gtc gtc ttg atc   | cgc atc aat gat ggg ggt cgg cca   | 2376 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| Glu Arg Glu Tyr Val Val Leu Ile   | Arg Ile Asn Asp Gly Gly Arg Pro   |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| 740                               | 745 750                           |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| ccc ttg gaa ggc att gtt tct tta   | cca gtt aca ttc tgc agt tgt gtg   | 2424 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| Pro Leu Glu Gly Ile Val Ser Leu   | Pro Val Thr Phe Cys Ser Cys Val   |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| 755                               | 760 765                           |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| gaa gga agt tgt ttc cgg cca gca   | ggt cac cag act ggg ata ccc act   | 2472 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| Glu Gly Ser Cys Phe Arg Pro Ala   | Gly His Gln Thr Gly Ile Pro Thr   |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| 770                               | 775 780                           |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| gtg ggc atg gca gtt ggt ata ctg   | ctg acc acc ctt ctg gtg att ggt   | 2520 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| Val Gly Met Ala Val Gly Ile Leu   | Leu Thr Thr Leu Leu Val Ile Gly   |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| 785                               | 790 795 800                       |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| ata att tta gca gtt gtg ttt atc   | cgc ata aag aag gat aaa ggc aaa   | 2568 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| Ile Ile Leu Ala Val Val Phe Ile   | Arg Ile Lys Lys Asp Lys Gly Lys   |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| 805                               | 810 815                           |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| gat aat gtt gaa agt gct caa gca   | tct gaa gtc aaa cct ctg aga agc   | 2616 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| Asp Asn Val Glu Ser Ala Gln Ala   | Ser Glu Val Lys Pro Leu Arg Ser   |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| 820                               | 825 830                           |      |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| tgaatttgaa aaggaatggt tgaatttata  | tagcaagtgc tatttcagca acaaccatct  | 2676 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| catcctatta cttttcatct aacgtgcatt  | ataatttttt aaacagatat tccctcttgt  | 2736 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| cctttaatat ttgctaaata tttctttttt  | gaggtggagt cttgctctgt cgcccaggct  | 2796 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| ggagtacagt ggtgtgatcc cagctcactg  | caacctccgc ctccctgggtt cacatgattc | 2856 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| tcctgcctca gcttcctaag tagctgggtt  | tacaggcacc caccaccatg cccagcta    | 2916 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| ttttgtatatt ttaatagaga cgggggtttc | gcatttggcc aggctgggtct tgaactcctg | 2976 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| acgtcaagtg atctgcctgc cttgggtctcc | caatacaggc atgaaccact gcacccacct  | 3036 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| acttagatat ttcattgtgt atagacatta  | gagagatttt tcatttttcc atgacatttt  | 3096 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| tcctctctgc aaatggctta gctacttgtg  | tttttccctt ttggggcaag acagactcat  | 3156 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| taaattattct gtacattttt tctttatcaa | ggagatatat cagtgttgct tcatagaact  | 3216 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
| gcctggattc catttatggt ttttctgatt  | ccatcctgtg tccccttcat ccttgactcc  | 3276 |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |

tttggtatct cactgaattt caaacatttg tcagagaaga aaaacgtgag gactcaggaa 3336  
 aaataaataa ataaaagaac agccttttcc cttagtatta acagaaatgt ttctgtgtca 3396  
 ttaaccatct ttaatcaatg tgacatgttg ctctttggct gaaattcttc aacttggaag 3456  
 tgacacagac ccacagaagg tgttcaaaca caacctactc tgcaaaccctt ggtaaaggaa 3516  
 ccagtcagct ggccagattt cctcactacc tgccatgcat acatgctgcy catgttttct 3576  
 tcattcgtat gttagtaaag ttttggttat tataatatta acatgtggaa gaaaacaaga 3636  
 catgaaaaga gtggtgacaa atcaagaata aacactgggt gtagtcagtt ttgtttgtta 3696  
 a 3797

<210> 60

<211> 832

<212> PRT

<213> NM\_004063 cadherin 17, CDH17

<400> 60

Met Ile Leu Gln Ala His Leu His Ser Leu Cys Leu Leu Met Leu Tyr  
 1 5 10 15

Leu Ala Thr Gly Tyr Gly Gln Glu Gly Lys Phe Ser Gly Pro Leu Lys  
 20 25 30

Pro Met Thr Phe Ser Ile Tyr Glu Gly Gln Glu Pro Ser Gln Ile Ile  
 35 40 45

Phe Gln Phe Lys Ala Asn Pro Pro Ala Val Thr Phe Glu Leu Thr Gly  
 50 55 60

Glu Thr Asp Asn Ile Phe Val Ile Glu Arg Glu Gly Leu Leu Tyr Tyr  
 65 70 75 80

Asn Arg Ala Leu Asp Arg Glu Thr Arg Ser Thr His Asn Leu Gln Val  
 85 90 95

Ala Ala Leu Asp Ala Asn Gly Ile Ile Val Glu Gly Pro Val Pro Ile  
 100 105 110

Thr Ile Glu Val Lys Asp Ile Asn Asp Asn Arg Pro Thr Phe Leu Gln  
 115 120 125

Ser Lys Tyr Glu Gly Ser Val Arg Gln Asn Ser Arg Pro Gly Lys Pro  
 130 135 140

Phe Leu Tyr Val Asn Ala Thr Asp Leu Asp Asp Pro Ala Thr Pro Asn  
 145 150 155 160

Gly Gln Leu Tyr Tyr Gln Ile Val Ile Gln Leu Pro Met Ile Asn Asn  
165 170 175

Val Met Tyr Phe Gln Ile Asn Asn Lys Thr Gly Ala Ile Ser Leu Thr  
180 185 190

Arg Glu Gly Ser Gln Glu Leu Asn Pro Ala Lys Asn Pro Ser Tyr Asn  
195 200 205

Leu Val Ile Ser Val Lys Asp Met Gly Gly Gln Ser Glu Asn Ser Phe  
210 215 220

Ser Asp Thr Thr Ser Val Asp Ile Ile Val Thr Glu Asn Ile Trp Lys  
225 230 235 240

Ala Pro Lys Pro Val Glu Met Val Glu Asn Ser Thr Asp Pro His Pro  
245 250 255

Ile Lys Ile Thr Gln Val Arg Trp Asn Asp Pro Gly Ala Gln Tyr Ser  
260 265 270

Leu Val Asp Lys Glu Lys Leu Pro Arg Phe Pro Phe Ser Ile Asp Gln  
275 280 285

Glu Gly Asp Ile Tyr Val Thr Gln Pro Leu Asp Arg Glu Glu Lys Asp  
290 295 300

Ala Tyr Val Phe Tyr Ala Val Ala Lys Asp Glu Tyr Gly Lys Pro Leu  
305 310 315 320

Ser Tyr Pro Leu Glu Ile His Val Lys Val Lys Asp Ile Asn Asp Asn  
325 330 335

Pro Pro Thr Cys Pro Ser Pro Val Thr Val Phe Glu Val Gln Glu Asn  
340 345 350

Glu Arg Leu Gly Asn Ser Ile Gly Thr Leu Thr Ala His Asp Arg Asp  
355 360 365

Glu Glu Asn Thr Ala Asn Ser Phe Leu Asn Tyr Arg Ile Val Glu Gln  
370 375 380

Thr Pro Lys Leu Pro Met Asp Gly Leu Phe Leu Ile Gln Thr Tyr Ala  
385 390 395 400

Gly Met Leu Gln Leu Ala Lys Gln Ser Leu Lys Lys Gln Asp Thr Pro  
405 410 415

Gln Tyr Asn Leu Thr Ile Glu Val Ser Asp Lys Asp Phe Lys Thr Leu  
420 425 430

Cys Phe Val Gln Ile Asn Val Ile Asp Ile Asn Asp Gln Ile Pro Ile  
435 440 445

Phe Glu Lys Ser Asp Tyr Gly Asn Leu Thr Leu Ala Glu Asp Thr Asn  
450 455 460

Ile Gly Ser Thr Ile Leu Thr Ile Gln Ala Thr Asp Ala Asp Glu Pro  
465 470 475 480

Phe Thr Gly Ser Ser Lys Ile Leu Tyr His Ile Ile Lys Gly Asp Ser  
485 490 495

Glu Gly Arg Leu Gly Val Asp Thr Asp Pro His Thr Asn Thr Gly Tyr  
500 505 510

Val Ile Ile Lys Lys Pro Leu Asp Phe Glu Thr Ala Ala Val Ser Asn  
515 520 525

Ile Val Phe Lys Ala Glu Asn Pro Glu Pro Leu Val Phe Gly Val Lys  
530 535 540

Tyr Asn Ala Ser Ser Phe Ala Lys Phe Thr Leu Ile Val Thr Asp Val  
545 550 555 560

Asn Glu Ala Pro Gln Phe Ser Gln His Val Phe Gln Ala Lys Val Ser  
565 570 575

Glu Asp Val Ala Ile Gly Thr Lys Val Gly Asn Val Thr Ala Lys Asp  
580 585 590

Pro Glu Gly Leu Asp Ile Ser Tyr Ser Leu Arg Gly Asp Thr Arg Gly  
595 600 605

Trp Leu Lys Ile Asp His Val Thr Gly Glu Ile Phe Ser Val Ala Pro  
610 615 620

Leu Asp Arg Glu Ala Gly Ser Pro Tyr Arg Val Gln Val Val Ala Thr  
625 630 635 640

Glu Val Gly Gly Ser Ser Leu Ser Ser Val Ser Glu Phe His Leu Ile  
645 650 655

Leu Met Asp Val Asn Asp Asn Pro Pro Arg Leu Ala Lys Asp Tyr Thr  
660 665 670

Gly Leu Phe Phe Cys His Pro Leu Ser Ala Pro Gly Ser Leu Ile Phe  
675 680 685

Glu Ala Thr Asp Asp Asp Gln His Leu Phe Arg Gly Pro His Phe Thr  
690 695 700

Phe Ser Leu Gly Ser Gly Ser Leu Gln Asn Asp Trp Glu Val Ser Lys  
705 710 715 720

Ile Asn Gly Thr His Ala Arg Leu Ser Thr Arg His Thr Glu Phe Glu  
725 730 735

Glu Arg Glu Tyr Val Val Leu Ile Arg Ile Asn Asp Gly Gly Arg Pro  
740 745 750

Pro Leu Glu Gly Ile Val Ser Leu Pro Val Thr Phe Cys Ser Cys Val  
755 760 765

Glu Gly Ser Cys Phe Arg Pro Ala Gly His Gln Thr Gly Ile Pro Thr  
770 775 780

Val Gly Met Ala Val Gly Ile Leu Leu Thr Thr Leu Leu Val Ile Gly  
785 790 795 800

Ile Ile Leu Ala Val Val Phe Ile Arg Ile Lys Lys Asp Lys Gly Lys  
805 810 815

Asp Asn Val Glu Ser Ala Gln Ala Ser Glu Val Lys Pro Leu Arg Ser  
820 825 830

<210> 61

<211> 2920

<212> DNA

<213> NM\_005588; meprin A, alpha

<220>

<221> CDS

<222> (10)..(2247)

<223>

<400> 61

cttgagca atg gct tgg att aga tcc act tgc att ctc ttt ttt acc ttg 51  
Met Ala Trp Ile Arg Ser Thr Cys Ile Leu Phe Phe Thr Leu  
1 5 10

ctt ttt gcc cac ata gca gct gta ccg att aag cat ctt cct gaa gaa 99  
Leu Phe Ala His Ile Ala Ala Val Pro Ile Lys His Leu Pro Glu Glu  
15 20 25 30

aat gta cat gat gca gat ttt ggt gaa cag aag gat att tca gaa atc 147  
Asn Val His Asp Ala Asp Phe Gly Glu Gln Lys Asp Ile Ser Glu Ile  
35 40 45

aat tta gct gca ggc ttg gac ctc ttt caa ggg gac atc ctc ttg cag 195  
Asn Leu Ala Ala Gly Leu Asp Leu Phe Gln Gly Asp Ile Leu Leu Gln

| 50  | 55  | 60  |      |
|---|-----|-----|------|
| aaa tcc aga aat ggc ctg aga gac cca aac acc agg tgg acg ttc ccc |     |     | 243  |
| Lys Ser Arg Asn Gly Leu Arg Asp Pro Asn Thr Arg Trp Thr Phe Pro |     |     |      |
| 65  | 70  | 75  |      |
| att cct tac atc ttg gct gat aat ttg ggg ctg aat gct aaa gga gcc |     |     | 291  |
| Ile Pro Tyr Ile Leu Ala Asp Asn Leu Gly Leu Asn Ala Lys Gly Ala |     |     |      |
| 80  | 85  | 90  |      |
| att ctg tat gcc ttt gag atg ttc cgt ctc aag tcc tgt gtg gat ttc |     |     | 339  |
| Ile Leu Tyr Ala Phe Glu Met Phe Arg Leu Lys Ser Cys Val Asp Phe |     |     |      |
| 95  | 100 | 105 | 110  |
| aag ccc tat gaa gga gag agc tca tat atc ata ttt caa cag ttt gat |     |     | 387  |
| Lys Pro Tyr Glu Gly Glu Ser Ser Tyr Ile Ile Phe Gln Gln Phe Asp |     |     |      |
| 115   | 120 | 125 |      |
| ggg tgc tgg tct gag gtt ggt gac caa cat gtg gga cag aac att tcc |     |     | 435  |
| Gly Cys Trp Ser Glu Val Gly Asp Gln His Val Gly Gln Asn Ile Ser |     |     |      |
| 130   | 135 | 140 |      |
| att ggc caa gga tgt gcc tat aag gcc atc ata gaa cac gag atc ctg |     |     | 483  |
| Ile Gly Gln Gly Cys Ala Tyr Lys Ala Ile Ile Glu His Glu Ile Leu |     |     |      |
| 145   | 150 | 155 |      |
| cat gct ttg gga ttt tac cac gag cag tca agg acg gac cgg gat gat |     |     | 531  |
| His Ala Leu Gly Phe Tyr His Glu Gln Ser Arg Thr Asp Arg Asp Asp |     |     |      |
| 160   | 165 | 170 |      |
| tat gtg aac atc tgg tgg gac caa att ctt tca ggt tac cag cac aac |     |     | 579  |
| Tyr Val Asn Ile Trp Trp Asp Gln Ile Leu Ser Gly Tyr Gln His Asn |     |     |      |
| 175   | 180 | 185 | 190  |
| ttt gac acc tat gat gat agc tta atc aca gac ctc aat aca ccc tat |     |     | 627  |
| Phe Asp Thr Tyr Asp Asp Ser Leu Ile Thr Asp Leu Asn Thr Pro Tyr |     |     |      |
| 195   | 200 | 205 |      |
| gat tat gag tct ttg atg cac tac cag cct ttc tca ttt aac aag aat |     |     | 675  |
| Asp Tyr Glu Ser Leu Met His Tyr Gln Pro Phe Ser Phe Asn Lys Asn |     |     |      |
| 210   | 215 | 220 |      |
| gca agt gtt ccc acc atc aca gcc aag atc cct gag ttt aac tcc att |     |     | 723  |
| Ala Ser Val Pro Thr Ile Thr Ala Lys Ile Pro Glu Phe Asn Ser Ile |     |     |      |
| 225   | 230 | 235 |      |
| atc gga caa cgc ctg gat ttc agt gcc att gat tta gag agg ctg aac |     |     | 771  |
| Ile Gly Gln Arg Leu Asp Phe Ser Ala Ile Asp Leu Glu Arg Leu Asn |     |     |      |
| 240   | 245 | 250 |      |
| cga atg tac aat tgc acc aca act cac act ctt ttg gac cac tgt act |     |     | 819  |
| Arg Met Tyr Asn Cys Thr Thr Thr His Thr Leu Leu Asp His Cys Thr |     |     |      |
| 255   | 260 | 265 | 270  |
| ttt gag aag gca aac atc tgt gga atg att cag ggc acc aga gat gac |     |     | 867  |
| Phe Glu Lys Ala Asn Ile Cys Gly Met Ile Gln Gly Thr Arg Asp Asp |     |     |      |
| 275   | 280 | 285 |      |
| act gac tgg gcc cat cag gac agt gct cag gct gga gaa gtg gat cac |     |     | 915  |
| Thr Asp Trp Ala His Gln Asp Ser Ala Gln Ala Gly Glu Val Asp His |     |     |      |
| 290   | 295 | 300 |      |
| acc ttg ttg gga caa tgc aca ggt gcc ggc tac ttc atg cag ttc agc |     |     | 963  |
| Thr Leu Leu Gly Gln Cys Thr Gly Ala Gly Tyr Phe Met Gln Phe Ser |     |     |      |
| 305   | 310 | 315 |      |
| acc agc tcg ggg tcc gcg gaa gag gca gcc cta ctg gag tct cgg att |     |     | 1011 |
| Thr Ser Ser Gly Ser Ala Glu Glu Ala Ala Leu Leu Glu Ser Arg Ile |     |     |      |



155

| 320   | 325 | 330 |      |
|---|-----|-----|------|
| ctt tac cca aag agg aag cag cag tgc ctg caa ttt ttc tat aaa atg<br>Leu Tyr Pro Lys Arg Lys Gln Gln Cys Leu Gln Phe Phe Tyr Lys Met<br>335 340 345 350 |     |     | 1059 |
| acg gga agt cct tca gac aga ctc gtt gtc tgg gtc agg agg gat gac<br>Thr Gly Ser Pro Ser Asp Arg Leu Val Val Trp Val Arg Arg Asp Asp<br>355 360 365     |     |     | 1107 |
| agc aca ggc aat gtt cgc aag ttg gtg aag gtg cag act ttt caa gga<br>Ser Thr Gly Asn Val Arg Lys Leu Val Lys Val Gln Thr Phe Gln Gly<br>370 375 380     |     |     | 1155 |
| gat gat gac cac aat tgg aaa att gcc cat gtg gtg ctc aaa gag gaa<br>Asp Asp Asp His Asn Trp Lys Ile Ala His Val Val Leu Lys Glu Glu<br>385 390 395     |     |     | 1203 |
| cag aag ttt cgc tac ctt ttc cag ggc aca aaa ggc gac cct cag aac<br>Gln Lys Phe Arg Tyr Leu Phe Gln Gly Thr Lys Gly Asp Pro Gln Asn<br>400 405 410     |     |     | 1251 |
| tca act ggg gga att tac cta gat gac atc act ctg aca gaa acc ccc<br>Ser Thr Gly Gly Ile Tyr Leu Asp Asp Ile Thr Leu Thr Glu Thr Pro<br>415 420 425 430 |     |     | 1299 |
| tgc ccc aca ggg gtc tgg aca gtc cgg aat ttc tcc caa gtc ctt gag<br>Cys Pro Thr Gly Val Trp Thr Val Arg Asn Phe Ser Gln Val Leu Glu<br>435 440 445     |     |     | 1347 |
| aac acc agc aaa ggg gac aag ctt cag agc cct cga ttc tac aat tgg<br>Asn Thr Ser Lys Gly Asp Lys Leu Gln Ser Pro Arg Phe Tyr Asn Ser<br>450 455 460     |     |     | 1395 |
| gag gga tat ggt ttt ggg gta act tta tac cca aat agc aga gaa agc<br>Glu Gly Tyr Gly Phe Gly Val Thr Leu Tyr Pro Asn Ser Arg Glu Ser<br>465 470 475     |     |     | 1443 |
| tct ggt tac ttg aga ctt gct ttt cat gtg tgc agt ggg gag aac gat<br>Ser Gly Tyr Leu Arg Leu Ala Phe His Val Cys Ser Gly Glu Asn Asp<br>480 485 490     |     |     | 1491 |
| gct atc ctg gag tgg ccg gta gaa aac aga cag gtg ata att acc atc<br>Ala Ile Leu Glu Trp Pro Val Glu Asn Arg Gln Val Ile Ile Thr Ile<br>495 500 505 510 |     |     | 1539 |
| ctt gac cag gag cct gat gtc cgg aac agg atg tcc tca agc atg gtg<br>Leu Asp Gln Glu Pro Asp Val Arg Asn Arg Met Ser Ser Ser Met Val<br>515 520 525     |     |     | 1587 |
| ttc act acc tcg aag tcg cac aca tct cca gcg ata aat gac act gtc<br>Phe Thr Thr Ser Lys Ser His Thr Ser Pro Ala Ile Asn Asp Thr Val<br>530 535 540     |     |     | 1635 |
| atc tgg gac agg ccg tcc agg gtg gga acc tat cat aca gac tgt aat<br>Ile Trp Asp Arg Pro Ser Arg Val Gly Thr Tyr His Thr Asp Cys Asn<br>545 550 555     |     |     | 1683 |
| tgt ttt aga agc atc gac ttg ggc tgg agt ggt ttc att tcc cac caa<br>Cys Phe Arg Ser Ile Asp Leu Gly Trp Ser Gly Phe Ile Ser His Gln<br>560 565 570     |     |     | 1731 |
| atg ctg aaa agg agg agt ttc ctg aaa aat gat gac ctc atc ata ttt<br>Met Leu Lys Arg Arg Ser Phe Leu Lys Asn Asp Asp Leu Ile Ile Phe<br>575 580 585 590 |     |     | 1779 |
| gtg gac ttt gaa gat atc acc cac ctc agc cag act gaa gtt ccc tct<br>Val Asp Phe Glu Asp Ile Thr His Leu Ser Gln Thr Glu Val Pro Ser                    |     |     | 1827 |

| 595   | 600 | 605 |      |
|---|-----|-----|------|
| aaa ggc aaa aga ctg agc ccc caa ggc ctc att ctc caa ggc cag gag<br>Lys Gly Lys Arg Leu Ser Pro Gln Gly Leu Ile Leu Gln Gly Gln Glu<br>610 615 620     |     |     | 1875 |
| cag cag gtc tcc gaa gaa ggt tgc gga aag gcc atg tta gag gaa gcc<br>Gln Gln Val Ser Glu Glu Gly Ser Gly Lys Ala Met Leu Glu Glu Ala<br>625 630 635     |     |     | 1923 |
| cta cct gtc agc ctg agc cag ggg cag ccc agc cga cag aag cgg tgc<br>Leu Pro Val Ser Leu Ser Pro Gln Gly Gln Pro Ser Arg Gln Lys Arg Ser<br>640 645 650 |     |     | 1971 |
| gtg gag aac aca ggc ccc ctg gag gac cat aac tgg cca cag tac ttc<br>Val Glu Asn Thr Gly Pro Leu Glu Asp His Asn Trp Pro Gln Tyr Phe<br>655 660 665 670 |     |     | 2019 |
| aga gac cca tgt gac cca aac cct tgc caa aat gac ggc atc tgt gtg<br>Arg Asp Pro Cys Asp Pro Asn Pro Cys Gln Asn Asp Gly Ile Cys Val<br>675 680 685     |     |     | 2067 |
| aac gtg aag ggg atg gcg agc tgc agg tgc atc tct gga cat gct ttc<br>Asn Val Lys Gly Met Ala Ser Cys Arg Cys Ile Ser Gly His Ala Phe<br>690 695 700     |     |     | 2115 |
| ttc tac acg ggg gag cgc tgt cag tgc gcc gag gtg cac ggc agt gtc<br>Phe Tyr Thr Gly Glu Arg Cys Gln Ser Ala Glu Val His Gly Ser Val<br>705 710 715     |     |     | 2163 |
| ctg ggc atg gtg atc gga ggc acg gct ggc gtg atc ttc ttg acc ttc<br>Leu Gly Met Val Ile Gly Gly Thr Ala Gly Val Ile Phe Leu Thr Phe<br>720 725 730     |     |     | 2211 |
| tcc atc atc gcc atc ctt tcc caa agg cca agg aag tgacctgcct<br>Ser Ile Ile Ala Ile Leu Ser Gln Arg Pro Arg Lys<br>735 740 745                          |     |     | 2257 |
| gctggcattg gccagaccac agcagcacct cctccatgca ggccttaact ttcccatgtt   |     |     | 2317 |
| caatgcagtt tggggcagct tttttatcag ccttgctttg gataggacct ccaaggacta   |     |     | 2377 |
| agcctccagc cccatgtgtg acccttgctca tctctctgcc ccacataatt atgttacttt  |     |     | 2437 |
| gctatgtgct cctaattgtat ctagtgtgtc ctgtgacaac actcatcaca cttcattgta  |     |     | 2497 |
| aatcacttgt tttattgact gtctttccta tagactgtaa gctccatgag ggcaggcaca   |     |     | 2557 |
| tggtgttctc attgaccgtg ctggccccag tgcctagatg catggctggc acattgttgg   |     |     | 2617 |
| cactcaacaa tgggtgaatg aataaaacaa taaatgaatg aataactaag atatagaac  |     |     | 2677 |
| tctcatttat attgcagatt gaatatatat gatgaaattc ttatgttgaa tatgttagaa   |     |     | 2737 |
| tcaaatactc atttttcatt agatacagta gtgtcatcac tcttttaaga tcttgttaaa   |     |     | 2797 |
| gatttcaaat aaaggtaact ctggcgagcc aggctgcaca gcatttgctt tcctctgaga   |     |     | 2857 |
| ttctaagaga aggcctttaa taaatttaaat aaatattgag ttagcaaaaa aaaaaaaaaa  |     |     | 2917 |
| aaa   |     |     | 2920 |

&lt;210&gt; 62

&lt;211&gt; 746

157

&lt;212&gt; PRT

&lt;213&gt; NM\_005588; meprin A, alpha

&lt;400&gt; 62

Met Ala Trp Ile Arg Ser Thr Cys Ile Leu Phe Phe Thr Leu Leu Phe  
1 5 10 15

Ala His Ile Ala Ala Val Pro Ile Lys His Leu Pro Glu Glu Asn Val  
20 25 30

His Asp Ala Asp Phe Gly Glu Gln Lys Asp Ile Ser Glu Ile Asn Leu  
35 40 45

Ala Ala Gly Leu Asp Leu Phe Gln Gly Asp Ile Leu Leu Gln Lys Ser  
50 55 60

Arg Asn Gly Leu Arg Asp Pro Asn Thr Arg Trp Thr Phe Pro Ile Pro  
65 70 75 80

Tyr Ile Leu Ala Asp Asn Leu Gly Leu Asn Ala Lys Gly Ala Ile Leu  
85 90 95

Tyr Ala Phe Glu Met Phe Arg Leu Lys Ser Cys Val Asp Phe Lys Pro  
100 105 110

Tyr Glu Gly Glu Ser Ser Tyr Ile Ile Phe Gln Gln Phe Asp Gly Cys  
115 120 125

Trp Ser Glu Val Gly Asp Gln His Val Gly Gln Asn Ile Ser Ile Gly  
130 135 140

Gln Gly Cys Ala Tyr Lys Ala Ile Ile Glu His Glu Ile Leu His Ala  
145 150 155 160

Leu Gly Phe Tyr His Glu Gln Ser Arg Thr Asp Arg Asp Asp Tyr Val  
165 170 175

Asn Ile Trp Trp Asp Gln Ile Leu Ser Gly Tyr Gln His Asn Phe Asp  
180 185 190

Thr Tyr Asp Asp Ser Leu Ile Thr Asp Leu Asn Thr Pro Tyr Asp Tyr  
195 200 205

Glu Ser Leu Met His Tyr Gln Pro Phe Ser Phe Asn Lys Asn Ala Ser  
210 215 220

Val Pro Thr Ile Thr Ala Lys Ile Pro Glu Phe Asn Ser Ile Ile Gly  
225 230 235 240

Gln Arg Leu Asp Phe Ser Ala Ile Asp Leu Glu Arg Leu Asn Arg Met  
245 250 255

Tyr Asn Cys Thr Thr Thr His Thr Leu Leu Asp His Cys Thr Phe Glu  
260 265 270

Lys Ala Asn Ile Cys Gly Met Ile Gln Gly Thr Arg Asp Asp Thr Asp  
275 280 285

Trp Ala His Gln Asp Ser Ala Gln Ala Gly Glu Val Asp His Thr Leu  
290 295 300

Leu Gly Gln Cys Thr Gly Ala Gly Tyr Phe Met Gln Phe Ser Thr Ser  
305 310 315 320

Ser Gly Ser Ala Glu Glu Ala Ala Leu Leu Glu Ser Arg Ile Leu Tyr  
325 330 335

Pro Lys Arg Lys Gln Gln Cys Leu Gln Phe Phe Tyr Lys Met Thr Gly  
340 345 350

Ser Pro Ser Asp Arg Leu Val Val Trp Val Arg Arg Asp Asp Ser Thr  
355 360 365

Gly Asn Val Arg Lys Leu Val Lys Val Gln Thr Phe Gln Gly Asp Asp  
370 375 380

Asp His Asn Trp Lys Ile Ala His Val Val Leu Lys Glu Glu Gln Lys  
385 390 395 400

Phe Arg Tyr Leu Phe Gln Gly Thr Lys Gly Asp Pro Gln Asn Ser Thr  
405 410 415

Gly Gly Ile Tyr Leu Asp Asp Ile Thr Leu Thr Glu Thr Pro Cys Pro  
420 425 430

Thr Gly Val Trp Thr Val Arg Asn Phe Ser Gln Val Leu Glu Asn Thr  
435 440 445

Ser Lys Gly Asp Lys Leu Gln Ser Pro Arg Phe Tyr Asn Ser Glu Gly  
450 455 460

Tyr Gly Phe Gly Val Thr Leu Tyr Pro Asn Ser Arg Glu Ser Ser Gly  
465 470 475 480

Tyr Leu Arg Leu Ala Phe His Val Cys Ser Gly Glu Asn Asp Ala Ile  
485 490 495

Leu Glu Trp Pro Val Glu Asn Arg Gln Val Ile Ile Thr Ile Leu Asp  
500 505 510

Gln Glu Pro Asp Val Arg Asn Arg Met Ser Ser Ser Met Val Phe Thr  
515 520 525

Thr Ser Lys Ser His Thr Ser Pro Ala Ile Asn Asp Thr Val Ile Trp  
530 535 540

Asp Arg Pro Ser Arg Val Gly Thr Tyr His Thr Asp Cys Asn Cys Phe  
545 550 555 560

Arg Ser Ile Asp Leu Gly Trp Ser Gly Phe Ile Ser His Gln Met Leu  
565 570 575

Lys Arg Arg Ser Phe Leu Lys Asn Asp Asp Leu Ile Ile Phe Val Asp  
580 585 590

Phe Glu Asp Ile Thr His Leu Ser Gln Thr Glu Val Pro Ser Lys Gly  
595 600 605

Lys Arg Leu Ser Pro Gln Gly Leu Ile Leu Gln Gly Gln Glu Gln Gln  
610 615 620

Val Ser Glu Glu Gly Ser Gly Lys Ala Met Leu Glu Glu Ala Leu Pro  
625 630 635 640

Val Ser Leu Ser Gln Gly Gln Pro Ser Arg Gln Lys Arg Ser Val Glu  
645 650 655

Asn Thr Gly Pro Leu Glu Asp His Asn Trp Pro Gln Tyr Phe Arg Asp  
660 665 670

Pro Cys Asp Pro Asn Pro Cys Gln Asn Asp Gly Ile Cys Val Asn Val  
675 680 685

Lys Gly Met Ala Ser Cys Arg Cys Ile Ser Gly His Ala Phe Phe Tyr  
690 695 700

Thr Gly Glu Arg Cys Gln Ser Ala Glu Val His Gly Ser Val Leu Gly  
705 710 715 720

Met Val Ile Gly Gly Thr Ala Gly Val Ile Phe Leu Thr Phe Ser Ile  
725 730 735

Ile Ala Ile Leu Ser Gln Arg Pro Arg Lys  
740 745

<210> 63

<211> 8838

<212> DNA

<213> NM\_015902 EDD, hyperplastic discs protein, DD5

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (34) .. (8430)

&lt;223&gt;

&lt;400&gt; 63

```

cgccctcgag tggaggacga gaaggaaagc acc atg acg tcc atc cat ttc gtg      54
               Met Thr Ser Ile His Phe Val
               1               5

ggt cac ccg ctg ccg ggc acc gag gac cag ctc aat gac agg tta cga      102
Val His Pro Leu Pro Gly Thr Glu Asp Gln Leu Asn Asp Arg Leu Arg
      10               15               20

gaa gtt tct gag aag ctg aac aaa tat aat tta aac agc cac ccc cct      150
Glu Val Ser Glu Lys Leu Asn Lys Tyr Asn Leu Asn Ser His Pro Pro
      25               30               35

ttg aat gta ttg gaa cag gct act att aaa cag tgt gtg gtg gga cca      198
Leu Asn Val Leu Glu Gln Ala Thr Ile Lys Gln Cys Val Val Gly Pro
      40               45               50               55

aat cat gct gcc ttt ctt ctt gag gat ggt aga gtt tgc agg att ggt      246
Asn His Ala Ala Phe Leu Leu Glu Asp Gly Arg Val Cys Arg Ile Gly
      60               65               70

ttt tca gta cag cca gac aga ttg gaa ttg ggt aaa cct gat aat aat      294
Phe Ser Val Gln Pro Asp Arg Leu Glu Leu Gly Lys Pro Asp Asn Asn
      75               80               85

gat ggg tca aag ttg aac agc aac tcg ggg gca ggg agg acg tca agg      342
Asp Gly Ser Lys Leu Asn Ser Asn Ser Gly Ala Gly Arg Thr Ser Arg
      90               95               100

cct ggt agg aca agc gac tct cca tgg ttt ctc tca ggt tct gag act      390
Pro Gly Arg Thr Ser Asp Ser Pro Trp Phe Leu Ser Gly Ser Glu Thr
      105               110               115

cta ggc agg ctg gca ggc aac acc tta gga agc cgc tgg agt tct gga      438
Leu Gly Arg Leu Ala Gly Asn Thr Leu Gly Ser Arg Trp Ser Ser Gly
      120               125               130               135

gtg ggt gga agt ggt gga gga tcc tct ggt agg tca tca gct gga gct      486
Val Gly Gly Ser Gly Gly Gly Ser Ser Gly Arg Ser Ser Ala Gly Ala
      140               145               150

cga gat tcc cgc cgg cag act cga gtt att cgg aca gga cgg gat cga      534
Arg Asp Ser Arg Arg Gln Thr Arg Val Ile Arg Thr Gly Arg Asp Arg
      155               160               165

ggg tct ggg ctt ttg ggc agt cag ccc cag cca gtt att cca gca tct      582
Gly Ser Gly Leu Leu Gly Ser Gln Pro Gln Pro Val Ile Pro Ala Ser
      170               175               180

gtc att cca gag gag ctg att tca cag gcc caa gtt gtt tta caa ggc      630
Val Ile Pro Glu Glu Leu Ile Ser Gln Ala Gln Val Val Leu Gln Gly
      185               190               195

aaa tcc aga agt gtc att att cga gaa ctt cag aga aca aat ctt gat      678
Lys Ser Arg Ser Val Ile Ile Arg Glu Leu Gln Arg Thr Asn Leu Asp
      200               205               210               215

```

|   |      |
|---|------|
| gtg aac ctt gct gta aat aat tta ctt agc cgg gat gat gaa gat gga<br>Val Asn Leu Ala Val Asn Asn Leu Leu Ser Arg Asp Asp Glu Asp Gly<br>220 225 230     | 726  |
| gat gat ggg gat gat aca gcc agc gaa tct tat ttg cct gga gag gat<br>Asp Asp Gly Asp Asp Thr Ala Ser Glu Ser Tyr Leu Pro Gly Glu Asp<br>235 240 245     | 774  |
| ctt atg tct ctc ctt gat gcc gac att cat tct gcc cac cca agt gtc<br>Leu Met Ser Leu Leu Asp Ala Asp Ile His Ser Ala His Pro Ser Val<br>250 255 260     | 822  |
| att att gat gca gat gcc atg ttt tct gaa gac att agc tat ttt ggt<br>Ile Ile Asp Ala Asp Ala Met Phe Ser Glu Asp Ile Ser Tyr Phe Gly<br>265 270 275     | 870  |
| tac cct tct ttt cgt cgt tca tca ctt tcc agg cta ggc tca tct cga<br>Tyr Pro Ser Phe Arg Arg Ser Ser Leu Ser Arg Leu Gly Ser Ser Arg<br>280 285 290 295 | 918  |
| gtt ctc ctt ctt ccc tta gag aga gac tct gag ctg ttg cgt gaa cgt<br>Val Leu Leu Leu Pro Leu Glu Arg Asp Ser Glu Leu Leu Arg Glu Arg<br>300 305 310     | 966  |
| gaa tcc gtt tta cgt tta cgt gaa cga agg tgg ctt gat gga gcc tca<br>Glu Ser Val Leu Arg Leu Arg Glu Arg Arg Trp Leu Asp Gly Ala Ser<br>315 320 325     | 1014 |
| ttt gat aat gaa agg ggt tct acc agc aag gaa gga gag cca aac ttg<br>Phe Asp Asn Glu Arg Gly Ser Thr Ser Lys Glu Gly Glu Pro Asn Leu<br>330 335 340     | 1062 |
| gat aag aag aat aca cct gtt caa agt cca gta tct cta gga gaa gat<br>Asp Lys Lys Asn Thr Pro Val Gln Ser Pro Val Ser Leu Gly Glu Asp<br>345 350 355     | 1110 |
| ttg cag tgg tgg cct gat aag gat gga aca aaa ttc atc tgt att ggg<br>Leu Gln Trp Trp Pro Asp Lys Asp Gly Thr Lys Phe Ile Cys Ile Gly<br>360 365 370 375 | 1158 |
| gct ctg tat tct gaa ctt ctg gct gtc agc agt aaa gga gaa ctt tat<br>Ala Leu Tyr Ser Glu Leu Leu Ala Val Ser Ser Lys Gly Glu Leu Tyr<br>380 385 390     | 1206 |
| cag tgg aaa tgg agt gaa tct gag cct tac aga aat gcc cag aat cct<br>Gln Trp Lys Trp Ser Glu Ser Glu Pro Tyr Arg Asn Ala Gln Asn Pro<br>395 400 405     | 1254 |
| tca tta cat cat cca cga gca aca ttt ttg ggg tta acc aat gaa aag<br>Ser Leu His His Pro Arg Ala Thr Phe Leu Gly Leu Thr Asn Glu Lys<br>410 415 420     | 1302 |
| ata gtc ctc ctg tct gca aat agc ata aga gca act gta gct aca gaa<br>Ile Val Leu Leu Ser Ala Asn Ser Ile Arg Ala Thr Val Ala Thr Glu<br>425 430 435     | 1350 |
| aat aac aag gtt gct aca tgg gtg gat gaa act tta agt tct gtg gct<br>Asn Asn Lys Val Ala Thr Trp Val Asp Glu Thr Leu Ser Ser Val Ala<br>440 445 450 455 | 1398 |
| tct aaa tta gag cac act gct cag act tac tct gaa ctt caa gga gag<br>Ser Lys Leu Glu His Thr Ala Gln Thr Tyr Ser Glu Leu Gln Gly Glu<br>460 465 470     | 1446 |
| cgg ata gtt tct tta cat tgc tgt gcc ctt tac acc tgc gct cag ctg<br>Arg Ile Val Ser Leu His Cys Cys Ala Leu Tyr Thr Cys Ala Gln Leu<br>475 480 485     | 1494 |

|   |      |
|---|------|
| gaa aac agt tta tat tgg tgg ggt gta gtt cct ttt agt caa agg aag<br>Glu Asn Ser Leu Tyr Trp Trp Gly Val Val Pro Phe Ser Gln Arg Lys<br>490 495 500     | 1542 |
| aaa atg tta gag aaa gct aga gca aaa aat aaa aag cct aaa tcc agt<br>Lys Met Leu Glu Lys Ala Arg Ala Lys Asn Lys Lys Pro Lys Ser Ser<br>505 510 515     | 1590 |
| gct ggt att tct tca atg ccg aac atc act gtt ggt acc cag gta tgc<br>Ala Gly Ile Ser Ser Met Pro Asn Ile Thr Val Gly Thr Gln Val Cys<br>520 525 530 535 | 1638 |
| ttg aga aat aat cct ctt tat cat gct gga gca gtt gca ttt tca att<br>Leu Arg Asn Asn Pro Leu Tyr His Ala Gly Ala Val Ala Phe Ser Ile<br>540 545 550     | 1686 |
| agt gct ggg att cct aaa gtt ggt gtc tta atg gag tca gtt tgg aat<br>Ser Ala Gly Ile Pro Lys Val Gly Val Leu Met Glu Ser Val Trp Asn<br>555 560 565     | 1734 |
| atg aat gac agc tgt aga ttt caa ctt aga tct cct gaa agc ttg aaa<br>Met Asn Asp Ser Cys Arg Phe Gln Leu Arg Ser Pro Glu Ser Leu Lys<br>570 575 580     | 1782 |
| aac atg gaa aaa gct agc aaa act act gaa gct aag cct gaa agt aag<br>Asn Met Glu Lys Ala Ser Lys Thr Thr Glu Ala Lys Pro Glu Ser Lys<br>585 590 595     | 1830 |
| cag gag cca gtg aaa aca gaa atg ggt cct cca cca tct cca gca tcc<br>Gln Glu Pro Val Lys Thr Glu Met Gly Pro Pro Ser Pro Ala Ser<br>600 605 610 615     | 1878 |
| acg tgt agt gat gca tcc tca att gcc agc agt gca tca atg cca tac<br>Thr Cys Ser Asp Ala Ser Ser Ile Ala Ser Ala Ser Met Pro Tyr<br>620 625 630         | 1926 |
| aaa cga cga cgg tca acc cct gca cca aaa gaa gag gaa aag gtg aat<br>Lys Arg Arg Arg Ser Thr Pro Ala Pro Lys Glu Glu Glu Lys Val Asn<br>635 640 645     | 1974 |
| gaa gag cag tgg tct ctt cgg gaa gtg gtt ttt gtg gaa gat gtc aag<br>Glu Glu Gln Trp Ser Leu Arg Glu Val Val Phe Val Glu Asp Val Lys<br>650 655 660     | 2022 |
| aat gtt cct gtt ggc aag gtg cta aaa gta gat ggt gcc tat gtt gct<br>Asn Val Pro Val Gly Lys Val Leu Lys Val Asp Gly Ala Tyr Val Ala<br>665 670 675     | 2070 |
| gta aaa ttt cca gga acc tcc agt aat act aac tgt cag aac agc tct<br>Val Lys Phe Pro Gly Thr Ser Ser Asn Thr Asn Cys Gln Asn Ser Ser<br>680 685 690 695 | 2118 |
| ggt cca gat gct gac cct tct tct ctc ctg cag gat tgt agg tta ctt<br>Gly Pro Asp Ala Asp Pro Ser Ser Leu Leu Gln Asp Cys Arg Leu Leu<br>700 705 710     | 2166 |
| aga att gat gaa ttg cag gtt gtc aaa act ggt gga aca ccg aag gtt<br>Arg Ile Asp Glu Leu Gln Val Val Lys Thr Gly Gly Thr Pro Lys Val<br>715 720 725     | 2214 |
| ccc gac tgt ttc caa agg act cct aaa aag ctt tgt ata cct gaa aaa<br>Pro Asp Cys Phe Gln Arg Thr Pro Lys Lys Leu Cys Ile Pro Glu Lys<br>730 735 740     | 2262 |
| aca gaa ata tta gca gtg aat gta gat tcc aaa ggt gtt cat gct gtt<br>Thr Glu Ile Leu Ala Val Asn Val Asp Ser Lys Gly Val His Ala Val<br>745 750 755     | 2310 |



|   |      |
|---|------|
| ctg aag act gga aat tgg gtg cga tac tgt atc ttt gat ctt gct aca<br>Leu Lys Thr Gly Asn Trp Val Arg Tyr Cys Ile Phe Asp Leu Ala Thr<br>760 765 770 775 | 2358 |
| gga aaa gca gaa cag gaa aat aat ttt cct aca agc agc att gct ttc<br>Gly Lys Ala Glu Gln Glu Asn Asn Phe Pro Thr Ser Ser Ile Ala Phe<br>780 785 790     | 2406 |
| ctt ggt cag aat gag agg aat gta gcc att ttc act gct gga cag gaa<br>Leu Gly Gln Asn Glu Arg Asn Val Ala Ile Phe Thr Ala Gly Gln Glu<br>795 800 805     | 2454 |
| tct ccc att att ctt cga gat gga aat ggt acc atc tac cca atg gcc<br>Ser Pro Ile Ile Leu Arg Asp Gly Asn Gly Thr Ile Tyr Pro Met Ala<br>810 815 820     | 2502 |
| aaa gat tgc atg gga gga ata agg gat ccc gat tgg ctg gat ctt cca<br>Lys Asp Cys Met Gly Gly Ile Arg Asp Pro Asp Trp Leu Asp Leu Pro<br>825 830 835     | 2550 |
| cct att agt agt ctt gga atg ggt gtg cat tct tta ata aat ctt cct<br>Pro Ile Ser Ser Leu Gly Met Gly Val His Ser Leu Ile Asn Leu Pro<br>840 845 850 855 | 2598 |
| gcc aat tca aca atc aaa aag aaa gct gct gtt atc atc atg gct gta<br>Ala Asn Ser Thr Ile Lys Lys Lys Ala Ala Val Ile Ile Met Ala Val<br>860 865 870     | 2646 |
| gag aaa caa acc tta atg caa cac att ctg cgc tgt gac tat gag gcc<br>Glu Lys Gln Thr Leu Met Gln His Ile Leu Arg Cys Asp Tyr Glu Ala<br>875 880 885     | 2694 |
| tgt cga caa tat cta atg aat ctt gag caa gcg gtt gtt tta gag cag<br>Cys Arg Gln Tyr Leu Met Asn Leu Glu Gln Ala Val Val Leu Glu Gln<br>890 895 900     | 2742 |
| aat cta cag atg ctg cag aca ttc atc agc cac aga tgt gat gga aat<br>Asn Leu Gln Met Leu Gln Thr Phe Ile Ser His Arg Cys Asp Gly Asn<br>905 910 915     | 2790 |
| cga aat att ttg cat gct tgt gta tca gtt tgc ttt cca acc agc aat<br>Arg Asn Ile Leu His Ala Cys Val Ser Val Cys Phe Pro Thr Ser Asn<br>920 925 930 935 | 2838 |
| aaa gaa act aaa gaa gaa gag gaa gcg gag cgt tct gaa aga aat aca<br>Lys Glu Thr Lys Glu Glu Glu Glu Ala Glu Arg Ser Glu Arg Asn Thr<br>940 945 950     | 2886 |
| ttt gca gaa agg ctt tct gct gtt gag gcc att gca aat gca ata tca<br>Phe Ala Glu Arg Leu Ser Ala Val Glu Ala Ile Ala Asn Ala Ile Ser<br>955 960 965     | 2934 |
| gtt gtt tca agt aat ggc cca ggt aat cgg gct gga tca tca agt agc<br>Val Val Ser Ser Asn Gly Pro Gly Asn Arg Ala Gly Ser Ser Ser Ser<br>970 975 980     | 2982 |
| cga agt ttg aga tta cgg gaa atg atg aga cgt tgc ttg aga gca gct<br>Arg Ser Leu Arg Leu Arg Glu Met Met Arg Arg Ser Leu Arg Ala Ala<br>985 990 995     | 3030 |
| ggt ttg ggt aga cat gaa gct gga gct tca tcc agt gac cac cag<br>Gly Leu Gly Arg His Glu Ala Gly Ala Ser Ser Ser Asp His Gln<br>1000 1005 1010          | 3075 |
| gat cca gtt tca ccc ccc ata gct ccc cct agt tgg gtt cct gac<br>Asp Pro Val Ser Pro Pro Ile Ala Pro Pro Ser Trp Val Pro Asp<br>1015 1020 1025          | 3120 |

|                    |  |  |  |                                    |      |
|--------------------|--|--|--|------------------------------------|------|
| cct<br>Pro<br>1030 | cct gcg atg gat<br>Pro Ala Met Asp         | cct<br>Pro<br>1035                         | gat ggt gac att gat<br>Asp Gly Asp Ile Asp | ttt atc ctg gcc<br>Phe Ile Leu Ala | 3165 |
| ccc<br>Pro<br>1045 | gct gtg gga tct<br>Ala Val Gly Ser         | ctt<br>Leu<br>1050                         | acc aca gca gca acc<br>Thr Thr Ala Ala Thr | ggt act ggt caa<br>Gly Thr Gly Gln | 3210 |
| gga<br>Gly<br>1060 | cca agc acc tcc<br>Pro Ser Thr Ser         | act<br>Thr<br>1065                         | att cca ggt cct tcc<br>Ile Pro Gly Pro Ser | aca gag cca tct<br>Thr Glu Pro Ser | 3255 |
| gta<br>Val<br>1075 | gta gaa tcc aag gat<br>Val Glu Ser Lys Asp | cga aag gcg aat gct<br>Arg Lys Ala Asn Ala | cat ttt ata ttg<br>His Phe Ile Leu         |                                    | 3300 |
| aaa<br>Lys<br>1090 | ttg tta tgt gac agt<br>Leu Leu Cys Asp Ser | gtg gtt ctc cag ccc<br>Val Val Leu Gln Pro | tat cta cga gaa<br>Tyr Leu Arg Glu         |                                    | 3345 |
| ctt<br>Leu<br>1105 | ctt tct gcc aag gat<br>Leu Ser Ala Lys Asp | gca aga ggg atg acc<br>Ala Arg Gly Met Thr | cca ttt atg tca<br>Pro Phe Met Ser         |                                    | 3390 |
| gct<br>Ala<br>1120 | gta agt ggc cga gct<br>Val Ser Gly Arg Ala | tat cct gct gca att<br>Tyr Pro Ala Ala Ile | acc atc tta gaa<br>Thr Ile Leu Glu         |                                    | 3435 |
| act<br>Thr<br>1135 | gct cag aaa att gca<br>Ala Gln Lys Ile Ala | aaa gct gaa ata tcc<br>Lys Ala Glu Ile Ser | tca agt gaa aaa<br>Ser Ser Glu Lys         |                                    | 3480 |
| gag<br>Glu<br>1150 | gaa gat gta ttc atg<br>Glu Asp Val Phe Met | gga atg gtt tgc cca<br>Gly Met Val Cys Pro | tca ggt acc aac<br>Ser Gly Thr Asn         |                                    | 3525 |
| cct<br>Pro<br>1165 | gat gac tct cct tta<br>Asp Asp Ser Pro Leu | tat gtt tta tgt tgt<br>Tyr Val Leu Cys Cys | aat gac act tgc<br>Asn Asp Thr Cys         |                                    | 3570 |
| agt<br>Ser<br>1180 | ttt aca tgg act gga<br>Phe Thr Trp Thr Gly | gca gag cac att aac<br>Ala Glu His Ile Asn | cag gat att ttt<br>Gln Asp Ile Phe         |                                    | 3615 |
| gag<br>Glu<br>1195 | tgt cga act tgt ggc<br>Cys Arg Thr Cys Gly | ttg ctg gag tca ctg<br>Leu Leu Glu Ser Leu | tgt tgt tgt acg<br>Cys Cys Cys Thr         |                                    | 3660 |
| gaa<br>Glu<br>1210 | tgt gca agg gtt tgt<br>Cys Ala Arg Val Cys | cat aaa ggt cat gat<br>His Lys Gly His Asp | tgc aaa ctc aaa<br>Cys Lys Leu Lys         |                                    | 3705 |
| cgg<br>Arg<br>1225 | aca tca cca aca gcc<br>Thr Ser Pro Thr Ala | tac tgt gat tgt tgg<br>Tyr Cys Asp Cys Trp | gag aaa tgt aaa<br>Glu Lys Cys Lys         |                                    | 3750 |
| tgt<br>Cys<br>1240 | aaa act ctt att gct<br>Lys Thr Leu Ile Ala | gga cag aaa tct gct<br>Gly Gln Lys Ser Ala | cgt ctt gat cta<br>Arg Leu Asp Leu         |                                    | 3795 |
| ctt<br>Leu<br>1255 | tat cgc ctg ctc act<br>Tyr Arg Leu Leu Thr | gct act aat ctg gtt<br>Ala Thr Asn Leu Val | act ctg cca aac<br>Thr Leu Pro Asn         |                                    | 3840 |
| agc<br>Ser<br>1270 | agg gga gag cac ctc<br>Arg Gly Glu His Leu | tta cta ttc tta gta<br>Leu Leu Phe Leu Val | cag aca gtc gca<br>Gln Thr Val Ala         |                                    | 3885 |

|                    |            |            |            |            |                    |            |            |            |            |                    |            |            |            |            |      |
|--------------------|------------|------------|------------|------------|--------------------|------------|------------|------------|------------|--------------------|------------|------------|------------|------------|------|
| agg<br>Arg<br>1285 | cag<br>Gln | acg<br>Thr | gtg<br>Val | gag<br>Glu | cat<br>His<br>1290 | tgt<br>Cys | caa<br>Gln | tac<br>Tyr | agg<br>Arg | cca<br>Pro<br>1295 | cct<br>Pro | cga<br>Arg | atc<br>Ile | agg<br>Arg | 3930 |
| gaa<br>Glu<br>1300 | gat<br>Asp | cgt<br>Arg | aac<br>Asn | cga<br>Arg | aaa<br>Lys<br>1305 | aca<br>Thr | gcc<br>Ala | agt<br>Ser | cct<br>Pro | gaa<br>Glu<br>1310 | gat<br>Asp | tca<br>Ser | gat<br>Asp | atg<br>Met | 3975 |
| cca<br>Pro<br>1315 | gat<br>Asp | cat<br>His | gat<br>Asp | tta<br>Leu | gag<br>Glu<br>1320 | cct<br>Pro | cca<br>Pro | aga<br>Arg | ttt<br>Phe | gcc<br>Ala<br>1325 | cag<br>Gln | ctt<br>Leu | gca<br>Ala | ttg<br>Leu | 4020 |
| gag<br>Glu<br>1330 | cgt<br>Arg | gtt<br>Val | cta<br>Leu | cag<br>Gln | gac<br>Asp<br>1335 | tgg<br>Trp | aat<br>Asn | gcc<br>Ala | ttg<br>Leu | aaa<br>Lys<br>1340 | tct<br>Ser | atg<br>Met | att<br>Ile | atg<br>Met | 4065 |
| ttt<br>Phe<br>1345 | ggg<br>Gly | tcg<br>Ser | cag<br>Gln | gag<br>Glu | aat<br>Asn<br>1350 | aaa<br>Lys | gac<br>Asp | cct<br>Pro | ctt<br>Leu | agt<br>Ser<br>1355 | gcc<br>Ala | agc<br>Ser | agt<br>Ser | aga<br>Arg | 4110 |
| ata<br>Ile<br>1360 | ggc<br>Gly | cat<br>His | ctt<br>Leu | ttg<br>Leu | cca<br>Pro<br>1365 | gaa<br>Glu | gag<br>Glu | caa<br>Gln | gta<br>Val | tac<br>Tyr<br>1370 | ctc<br>Leu | aat<br>Asn | cag<br>Gln | caa<br>Gln | 4155 |
| agt<br>Ser<br>1375 | ggc<br>Gly | aca<br>Thr | att<br>Ile | cgg<br>Arg | ctg<br>Leu<br>1380 | gac<br>Asp | tgt<br>Cys | ttc<br>Phe | act<br>Thr | cat<br>His<br>1385 | tgc<br>Cys | ctt<br>Leu | ata<br>Ile | gtt<br>Val | 4200 |
| aag<br>Lys<br>1390 | tgt<br>Cys | aca<br>Thr | gca<br>Ala | gat<br>Asp | att<br>Ile<br>1395 | ttg<br>Leu | ctt<br>Leu | tta<br>Leu | gat<br>Asp | act<br>Thr<br>1400 | cta<br>Leu | cta<br>Leu | ggt<br>Gly | aca<br>Thr | 4245 |
| cta<br>Leu<br>1405 | gtg<br>Val | aaa<br>Lys | gaa<br>Glu | ctc<br>Leu | caa<br>Gln<br>1410 | aac<br>Asn | aaa<br>Lys | tat<br>Tyr | aca<br>Thr | cct<br>Pro<br>1415 | gga<br>Gly | cgt<br>Arg | aga<br>Arg | gaa<br>Glu | 4290 |
| gaa<br>Glu<br>1420 | gct<br>Ala | att<br>Ile | gct<br>Ala | gtg<br>Val | aca<br>Thr<br>1425 | atg<br>Met | agg<br>Arg | ttt<br>Phe | cta<br>Leu | cgt<br>Arg<br>1430 | tca<br>Ser | gtg<br>Val | gca<br>Ala | aga<br>Arg | 4335 |
| gtt<br>Val<br>1435 | ttt<br>Phe | gtt<br>Val | att<br>Ile | ctg<br>Leu | agt<br>Ser<br>1440 | gtg<br>Val | gaa<br>Glu | atg<br>Met | gct<br>Ala | tca<br>Ser<br>1445 | tcc<br>Ser | aaa<br>Lys | aag<br>Lys | aaa<br>Lys | 4380 |
| aac<br>Asn<br>1450 | aac<br>Asn | ttt<br>Phe | att<br>Ile | cca<br>Pro | cag<br>Gln<br>1455 | cca<br>Pro | att<br>Ile | gga<br>Gly | aaa<br>Lys | tgc<br>Cys<br>1460 | aag<br>Lys | cgt<br>Arg | gta<br>Val | ttc<br>Phe | 4425 |
| caa<br>Gln<br>1465 | gca<br>Ala | ttg<br>Leu | cta<br>Leu | cct<br>Pro | tac<br>Tyr<br>1470 | gct<br>Ala | gtg<br>Val | gaa<br>Glu | gaa<br>Glu | ttg<br>Leu<br>1475 | tgc<br>Cys | aac<br>Asn | gta<br>Val | gca<br>Ala | 4470 |
| gag<br>Glu<br>1480 | tca<br>Ser | ctg<br>Leu | att<br>Ile | gtt<br>Val | cct<br>Pro<br>1485 | gtc<br>Val | aga<br>Arg | atg<br>Met | ggg<br>Gly | att<br>Ile<br>1490 | gct<br>Ala | cgt<br>Arg | cca<br>Pro | act<br>Thr | 4515 |
| gca<br>Ala<br>1495 | cca<br>Pro | ttt<br>Phe | acc<br>Thr | ctg<br>Leu | gct<br>Ala<br>1500 | agt<br>Ser | act<br>Thr | agc<br>Ser | ata<br>Ile | gat<br>Asp<br>1505 | gcc<br>Ala | atg<br>Met | cag<br>Gln | ggc<br>Gly | 4560 |
| agt<br>Ser<br>1510 | gaa<br>Glu | gaa<br>Glu | tta<br>Leu | ttt<br>Phe | tca<br>Ser<br>1515 | gtg<br>Val | gaa<br>Glu | cca<br>Pro | cta<br>Leu | cca<br>Pro<br>1520 | cca<br>Pro | cga<br>Arg | cca<br>Pro | tca<br>Ser | 4605 |
| tct<br>Ser<br>1525 | gat<br>Asp | cag<br>Gln | tct<br>Ser | agc<br>Ser | agc<br>Ser<br>1530 | tcc<br>Ser | agt<br>Ser | cag<br>Gln | tct<br>Ser | cag<br>Gln<br>1535 | tca<br>Ser | tcc<br>Ser | tac<br>Tyr | atc<br>Ile | 4650 |

166

|                    |            |            |            |            |            |            |            |            |            |            |            |            |            |            |      |
|--------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------|
| atc<br>Ile<br>1540 | agg<br>Arg | aat<br>Asn | cca<br>Pro | cag<br>Gln | cag<br>Gln | agg<br>Arg | cgc<br>Arg | atc<br>Ile | agc<br>Ser | cag<br>Gln | tca<br>Ser | cag<br>Gln | ccc<br>Pro | gtt<br>Val | 4695 |
|                    |            |            |            |            | 1545       |            |            |            |            | 1550       |            |            |            |            |      |
| cgg<br>Arg<br>1555 | ggc<br>Gly | aga<br>Arg | gat<br>Asp | gaa<br>Glu | gaa<br>Glu | cag<br>Gln | gat<br>Asp | gat<br>Asp | att<br>Ile | gtt<br>Val | tca<br>Ser | gca<br>Ala | gat<br>Asp | gtg<br>Val | 4740 |
|                    |            |            |            |            | 1560       |            |            |            |            | 1565       |            |            |            |            |      |
| gaa<br>Glu<br>1570 | gag<br>Glu | gtt<br>Val | gag<br>Glu | gtg<br>Val | gtg<br>Val | gag<br>Glu | ggg<br>Gly | gtg<br>Val | gct<br>Ala | gga<br>Gly | gaa<br>Glu | gag<br>Glu | gat<br>Asp | cat<br>His | 4785 |
|                    |            |            |            |            | 1575       |            |            |            |            | 1580       |            |            |            |            |      |
| cat<br>His<br>1585 | gat<br>Asp | gaa<br>Glu | cag<br>Gln | gaa<br>Glu | gaa<br>Glu | cac<br>His | ggg<br>Gly | gaa<br>Glu | gaa<br>Glu | aat<br>Asn | gct<br>Ala | gag<br>Glu | gca<br>Ala | gag<br>Glu | 4830 |
|                    |            |            |            |            | 1590       |            |            |            |            | 1595       |            |            |            |            |      |
| gga<br>Gly<br>1600 | caa<br>Gln | cat<br>His | gat<br>Asp | gag<br>Glu | cat<br>His | gat<br>Asp | gaa<br>Glu | gac<br>Asp | ggg<br>Gly | agt<br>Ser | gat<br>Asp | atg<br>Met | gag<br>Glu | ctg<br>Leu | 4875 |
|                    |            |            |            |            | 1605       |            |            |            |            | 1610       |            |            |            |            |      |
| gac<br>Asp<br>1615 | ttg<br>Leu | tta<br>Leu | gca<br>Ala | gca<br>Ala | gct<br>Ala | gaa<br>Glu | aca<br>Thr | gaa<br>Glu | agt<br>Ser | gat<br>Asp | agt<br>Ser | gaa<br>Glu | agt<br>Ser | aac<br>Asn | 4920 |
|                    |            |            |            |            | 1620       |            |            |            |            | 1625       |            |            |            |            |      |
| cac<br>His<br>1630 | agc<br>Ser | aac<br>Asn | caa<br>Gln | gat<br>Asp | aat<br>Asn | gct<br>Ala | agt<br>Ser | ggg<br>Gly | cgc<br>Arg | aga<br>Arg | agc<br>Ser | gtt<br>Val | gtc<br>Val | act<br>Thr | 4965 |
|                    |            |            |            |            | 1635       |            |            |            |            | 1640       |            |            |            |            |      |
| gca<br>Ala<br>1645 | gca<br>Ala | act<br>Thr | gct<br>Ala | ggt<br>Gly | tca<br>Ser | gaa<br>Glu | gca<br>Ala | gga<br>Gly | gca<br>Ala | agc<br>Ser | agt<br>Ser | gtt<br>Val | cct<br>Pro | gcc<br>Ala | 5010 |
|                    |            |            |            |            | 1650       |            |            |            |            | 1655       |            |            |            |            |      |
| ttc<br>Phe<br>1660 | ttt<br>Phe | tct<br>Ser | gaa<br>Glu | gat<br>Asp | gat<br>Asp | tct<br>Ser | caa<br>Gln | tcg<br>Ser | aat<br>Asn | gac<br>Asp | tca<br>Ser | agt<br>Ser | gat<br>Ser | tct<br>Ser | 5055 |
|                    |            |            |            |            | 1665       |            |            |            |            | 1670       |            |            |            |            |      |
| gat<br>Asp<br>1675 | agc<br>Ser | agt<br>Ser | agt<br>Ser | agt<br>Ser | cag<br>Gln | agt<br>Ser | gac<br>Asp | gac<br>Asp | ata<br>Ile | gaa<br>Glu | cag<br>Gln | gag<br>Glu | acc<br>Thr | ttt<br>Phe | 5100 |
|                    |            |            |            |            | 1680       |            |            |            |            | 1685       |            |            |            |            |      |
| atg<br>Met<br>1690 | ctt<br>Leu | gat<br>Asp | gag<br>Glu | cca<br>Pro | tta<br>Leu | gaa<br>Glu | aga<br>Arg | acc<br>Thr | aca<br>Thr | aat<br>Asn | agc<br>Ser | tcc<br>Ser | cat<br>His | gcc<br>Ala | 5145 |
|                    |            |            |            |            | 1695       |            |            |            |            | 1700       |            |            |            |            |      |
| aat<br>Asn<br>1705 | ggt<br>Gly | gct<br>Ala | gcc<br>Ala | caa<br>Gln | gct<br>Ala | ccc<br>Pro | cgt<br>Arg | tca<br>Ser | atg<br>Met | cag<br>Gln | tgg<br>Trp | gct<br>Ala | gtc<br>Val | cgc<br>Arg | 5190 |
|                    |            |            |            |            | 1710       |            |            |            |            | 1715       |            |            |            |            |      |
| aac<br>Asn<br>1720 | acc<br>Thr | cag<br>Gln | cat<br>His | cag<br>Gln | cga<br>Arg | gca<br>Ala | gcc<br>Ala | agt<br>Ser | aca<br>Thr | gcc<br>Ala | cct<br>Pro | tcc<br>Ser | agt<br>Ser | aca<br>Thr | 5235 |
|                    |            |            |            |            | 1725       |            |            |            |            | 1730       |            |            |            |            |      |
| tct<br>Ser<br>1735 | aca<br>Thr | cca<br>Pro | gca<br>Ala | gca<br>Ala | agt<br>Ser | tca<br>Ser | gcg<br>Ala | ggg<br>Gly | ttg<br>Leu | att<br>Ile | tat<br>Tyr | att<br>Ile | gat<br>Asp | cct<br>Pro | 5280 |
|                    |            |            |            |            | 1740       |            |            |            |            | 1745       |            |            |            |            |      |
| tca<br>Ser<br>1750 | aac<br>Asn | tta<br>Leu | cgc<br>Arg | cgg<br>Arg | agt<br>Ser | ggg<br>Gly | acc<br>Thr | atc<br>Ile | agt<br>Ser | aca<br>Thr | agt<br>Ser | gct<br>Ala | gca<br>Ala | gct<br>Ala | 5325 |
|                    |            |            |            |            | 1755       |            |            |            |            | 1760       |            |            |            |            |      |
| gca<br>Ala<br>1765 | gca<br>Ala | gct<br>Ala | gct<br>Ala | ttg<br>Leu | gaa<br>Glu | gct<br>Ala | agc<br>Ser | aac<br>Asn | gcc<br>Ala | agc<br>Ser | agt<br>Ser | tac<br>Tyr | cta<br>Leu | aca<br>Thr | 5370 |
|                    |            |            |            |            | 1770       |            |            |            |            | 1775       |            |            |            |            |      |
| tct<br>Ser<br>1780 | gca<br>Ala | agc<br>Ser | agt<br>Ser | tta<br>Leu | gcc<br>Ala | agg<br>Arg | gct<br>Ala | tac<br>Tyr | agc<br>Ser | att<br>Ile | gtc<br>Val | att<br>Ile | aga<br>Arg | caa<br>Gln | 5415 |
|                    |            |            |            |            | 1785       |            |            |            |            | 1790       |            |            |            |            |      |

|                    |  |  |                                    |      |
|--------------------|--|--|------------------------------------|------|
| atc<br>Ile<br>1795 | tcg gac ttg atg ggc<br>Ser Asp Leu Met Gly<br>1800 | ctt att cct aag tat<br>Leu Ile Pro Lys Tyr<br>1805 | aat cac cta gta<br>Asn His Leu Val | 5460 |
| tac<br>Tyr<br>1810 | tct cag att cca gca<br>Ser Gln Ile Pro Ala<br>1815 | gct gtg aaa ttg act<br>Ala Val Lys Leu Thr<br>1820 | tac caa gat gca<br>Tyr Gln Asp Ala | 5505 |
| gta<br>Val<br>1825 | aac tta cag aac tat<br>Asn Leu Gln Asn Tyr<br>1830 | gta gaa gaa aag ctt<br>Val Glu Glu Lys Leu<br>1835 | att ccc act tgg<br>Ile Pro Thr Trp | 5550 |
| aac<br>Asn<br>1840 | tgg atg gtc agt att<br>Trp Met Val Ser Ile<br>1845 | atg gat tct act gaa<br>Met Asp Ser Thr Glu<br>1850 | gct caa tta cgt<br>Ala Gln Leu Arg | 5595 |
| tat<br>Tyr<br>1855 | ggt tct gca tta gca<br>Gly Ser Ala Leu Ala<br>1860 | tct gct ggt gat cct<br>Ser Ala Gly Asp Pro<br>1865 | gga cat cca aat<br>Gly His Pro Asn | 5640 |
| cat<br>His<br>1870 | cct ctt cac gct tct<br>Pro Leu His Ala Ser<br>1875 | cag aat tca gcg aga<br>Gln Asn Ser Ala Arg<br>1880 | aga gag agg atg<br>Arg Glu Arg Met | 5685 |
| act<br>Thr<br>1885 | gcg cga gaa gaa gct<br>Ala Arg Glu Glu Ala<br>1890 | agc tta cga aca ctt<br>Ser Leu Arg Thr Leu<br>1895 | gaa ggc aga cga<br>Glu Gly Arg Arg | 5730 |
| cgt<br>Arg<br>1900 | gcc acc ttg ctt agc<br>Ala Thr Leu Leu Ser<br>1905 | gcc cgt caa gga atg<br>Ala Arg Gln Gly Met<br>1910 | atg tct gca cga<br>Met Ser Ala Arg | 5775 |
| gga<br>Gly<br>1915 | gac ttc cta aat tat<br>Asp Phe Leu Asn Tyr<br>1920 | gct ctg tct cta atg<br>Ala Leu Ser Leu Met<br>1925 | cgg tct cat aat<br>Arg Ser His Asn | 5820 |
| gat<br>Asp<br>1930 | gag cat tct gat gtt<br>Glu His Ser Asp Val<br>1935 | ctt cca gtt ttg gat<br>Leu Pro Val Leu Asp<br>1940 | gtt tgc tca ttg<br>Val Cys Ser Leu | 5865 |
| aag<br>Lys<br>1945 | cat gtg gca tat gtt<br>His Val Ala Tyr Val<br>1950 | ttt caa gca ctt ata<br>Phe Gln Ala Leu Ile<br>1955 | tac tgg att aag<br>Tyr Trp Ile Lys | 5910 |
| gca<br>Ala<br>1960 | atg aat cag cag aca<br>Met Asn Gln Gln Thr<br>1965 | aca ttg gat aca cct<br>Thr Leu Asp Thr Pro<br>1970 | caa cta gaa cgc<br>Gln Leu Glu Arg | 5955 |
| aaa<br>Lys<br>1975 | agg acg cga gaa ctc<br>Arg Thr Arg Glu Leu<br>1980 | ttg gaa ctg ggt att<br>Leu Glu Leu Gly Ile<br>1985 | gat aat gaa gat<br>Asp Asn Glu Asp | 6000 |
| tca<br>Ser<br>1990 | gaa cat gaa aat gat<br>Glu His Glu Asn Asp<br>1995 | gat gac acc aat caa<br>Asp Asp Thr Asn Gln<br>2000 | agt gct act ttg<br>Ser Ala Thr Leu | 6045 |
| aat<br>Asn<br>2005 | gat aag gat gat gac<br>Asp Lys Asp Asp Asp<br>2010 | tct ctt cct gca gaa<br>Ser Leu Pro Ala Glu<br>2015 | act ggc caa aac<br>Thr Gly Gln Asn | 6090 |
| cat<br>His<br>2020 | cca ttt ttc cga cgt<br>Pro Phe Phe Arg Arg<br>2025 | tca gac tcc atg aca<br>Ser Asp Ser Met Thr<br>2030 | ttc ctt ggg tgt<br>Phe Leu Gly Cys | 6135 |
| ata<br>Ile<br>2035 | ccc cca aat cca ttt<br>Pro Pro Asn Pro Phe<br>2040 | gaa gtg cct ctg gct<br>Glu Val Pro Leu Ala<br>2045 | gaa gcc atc ccc<br>Glu Ala Ile Pro | 6180 |

|                         |                     |                 |      |
|-------------------------|---------------------|-----------------|------|
| ttg gct gat cag cca cat | ctg ttg cag cca aat | gct aga aag gag | 6225 |
| Leu Ala Asp Gln Pro His | Leu Leu Gln Pro Asn | Ala Arg Lys Glu |      |
| 2050                    | 2055                | 2060            |      |
| gat ctt ttt ggc cgt cca | agt cag ggt ctt tat | tct tca tct gcc | 6270 |
| Asp Leu Phe Gly Arg Pro | Ser Gln Gly Leu Tyr | Ser Ser Ser Ala |      |
| 2065                    | 2070                | 2075            |      |
| agt agt ggg aaa tgt tta | atg gag gtt aca gtg | gat aga aac tgc | 6315 |
| Ser Ser Gly Lys Cys Leu | Met Glu Val Thr Val | Asp Arg Asn Cys |      |
| 2080                    | 2085                | 2090            |      |
| cta gag gtt ctt cca aca | aaa atg tct tat gct | gcc aat ctg aaa | 6360 |
| Leu Glu Val Leu Pro Thr | Lys Met Ser Tyr Ala | Ala Asn Leu Lys |      |
| 2095                    | 2100                | 2105            |      |
| aat gta atg aac atg caa | aac cgg caa aaa aaa | gaa ggg gaa gaa | 6405 |
| Asn Val Met Asn Met Gln | Asn Arg Gln Lys Lys | Glu Gly Glu Glu |      |
| 2110                    | 2115                | 2120            |      |
| cag ccc gtg ctg cca gaa | gaa act gag agt tca | aaa cca ggg cca | 6450 |
| Gln Pro Val Leu Pro Glu | Glu Thr Glu Ser Ser | Lys Pro Gly Pro |      |
| 2125                    | 2130                | 2135            |      |
| tct gct cat gat ctt gct | gca caa tta aaa agt | agc tta cta gca | 6495 |
| Ser Ala His Asp Leu Ala | Ala Gln Leu Lys Ser | Ser Leu Leu Ala |      |
| 2140                    | 2145                | 2150            |      |
| gaa ata gga ctt act gaa | agt gaa ggg cca cct | ctc aca tct ttc | 6540 |
| Glu Ile Gly Leu Thr Glu | Ser Glu Gly Pro Pro | Leu Thr Ser Phe |      |
| 2155                    | 2160                | 2165            |      |
| agg cca cag tgt agc ttt | atg gga atg gtt att | tcc cat gat atg | 6585 |
| Arg Pro Gln Cys Ser Phe | Met Gly Met Val Ile | Ser His Asp Met |      |
| 2170                    | 2175                | 2180            |      |
| ctg cta gga cgt tgg cgc | ctt tct tta gaa ctg | ttc ggc agg gta | 6630 |
| Leu Leu Gly Arg Trp Arg | Leu Ser Leu Glu Leu | Phe Gly Arg Val |      |
| 2185                    | 2190                | 2195            |      |
| ttc atg gaa gat gtt gga | gca gaa cct gga tca | atc cta act gaa | 6675 |
| Phe Met Glu Asp Val Gly | Ala Glu Pro Gly Ser | Ile Leu Thr Glu |      |
| 2200                    | 2205                | 2210            |      |
| ttg ggt ggt ttt gag gta | aaa gaa tca aaa ttc | cgc aga gaa atg | 6720 |
| Leu Gly Gly Phe Glu Val | Lys Glu Ser Lys Phe | Arg Arg Glu Met |      |
| 2215                    | 2220                | 2225            |      |
| gaa aaa ctg aga aac cag | cag tca aga gat ttg | tca cta gag gtt | 6765 |
| Glu Lys Leu Arg Asn Gln | Gln Ser Arg Asp Leu | Ser Leu Glu Val |      |
| 2230                    | 2235                | 2240            |      |
| gat cgg gat cga gat ctt | ctc att cag cag act | atg agg cag ctt | 6810 |
| Asp Arg Asp Arg Asp Leu | Leu Ile Gln Gln Thr | Met Arg Gln Leu |      |
| 2245                    | 2250                | 2255            |      |
| aac aat cac ttt ggt cga | aga tgt gct act aca | cca atg gct gta | 6855 |
| Asn Asn His Phe Gly Arg | Arg Cys Ala Thr Thr | Pro Met Ala Val |      |
| 2260                    | 2265                | 2270            |      |
| cac aga gta aaa gtc aca | ttt aag gat gag cca | gga gag ggc agt | 6900 |
| His Arg Val Lys Val Thr | Phe Lys Asp Glu Pro | Gly Glu Gly Ser |      |
| 2275                    | 2280                | 2285            |      |
| ggt gta gca cga agt ttt | tat aca gcc att gca | caa gca ttt tta | 6945 |
| Gly Val Ala Arg Ser Phe | Tyr Thr Ala Ile Ala | Gln Ala Phe Leu |      |
| 2290                    | 2295                | 2300            |      |

|                    |  |  |                                    |      |
|--------------------|--|--|------------------------------------|------|
| tca<br>Ser<br>2305 | aat gaa aaa ttg cca<br>Asn Glu Lys Leu Pro<br>2310 | aat cta gag tgt atc<br>Asn Leu Glu Cys Ile<br>2315 | caa aat gcc aac<br>Gln Asn Ala Asn | 6990 |
| aaa<br>Lys<br>2320 | ggc acc cac aca agt<br>Gly Thr His Thr Ser<br>2325 | tta atg cag aga tta<br>Leu Met Gln Arg Leu<br>2330 | agg aac cga gga<br>Arg Asn Arg Gly | 7035 |
| gag<br>Glu<br>2335 | aga gac cgg gaa agg<br>Arg Asp Arg Glu Arg<br>2340 | gag aga gaa agg gaa<br>Glu Arg Glu Arg Glu<br>2345 | atg agg agg agt<br>Met Arg Arg Ser | 7080 |
| agt<br>Ser<br>2350 | ggg ttg cga gca ggt<br>Gly Leu Arg Ala Gly<br>2355 | tct cgg agg gac cgg<br>Ser Arg Arg Asp Arg<br>2360 | gat aga gac ttt<br>Asp Arg Asp Phe | 7125 |
| aga<br>Arg<br>2365 | aga cag ctt tcc atc<br>Arg Gln Leu Ser Ile<br>2370 | gac act agg ccc ttt<br>Asp Thr Arg Pro Phe<br>2375 | aga cca gcc tct<br>Arg Pro Ala Ser | 7170 |
| gaa<br>Glu<br>2380 | ggg aat cct agc gat<br>Gly Asn Pro Ser Asp<br>2385 | gat cct gag cct ttg<br>Asp Pro Glu Pro Leu<br>2390 | cca gca cat cgg<br>Pro Ala His Arg | 7215 |
| cag<br>Gln<br>2395 | gca ctt gga gag agg<br>Ala Leu Gly Glu Arg<br>2400 | ctt tat cct cgt gta<br>Leu Tyr Pro Arg Val<br>2405 | caa gca atg caa<br>Gln Ala Met Gln | 7260 |
| cca<br>Pro<br>2410 | gca ttt gca agt aaa<br>Ala Phe Ala Ser Lys<br>2415 | atc act ggc atg ttg<br>Ile Thr Gly Met Leu<br>2420 | ttg gaa tta tcc<br>Leu Glu Leu Ser | 7305 |
| cca<br>Pro<br>2425 | gct cag ctg ctt ctc<br>Ala Gln Leu Leu Leu<br>2430 | ctt cta gca agt gag<br>Leu Leu Ala Ser Glu<br>2435 | gat tct ctg aga<br>Asp Ser Leu Arg | 7350 |
| gca<br>Ala<br>2440 | aga gtg gat gag gcc<br>Arg Val Asp Glu Ala<br>2445 | atg gaa ctc att att<br>Met Glu Leu Ile Ile<br>2450 | gca cat gga cgg<br>Ala His Gly Arg | 7395 |
| gaa<br>Glu<br>2455 | aat gga gct gat agt<br>Asn Gly Ala Asp Ser<br>2460 | atc ctg gat ctt gga<br>Ile Leu Asp Leu Gly<br>2465 | tta gta gac tcc<br>Leu Val Asp Ser | 7440 |
| tca<br>Ser<br>2470 | gaa aag gta cag cag<br>Glu Lys Val Gln Gln<br>2475 | gaa aac cga aag cgc<br>Glu Asn Arg Lys Arg<br>2480 | cat ggc tct agt<br>His Gly Ser Ser | 7485 |
| cga<br>Arg<br>2485 | agt gta gta gat atg<br>Ser Val Val Asp Met<br>2490 | gat tta gat gat aca<br>Asp Leu Asp Asp Thr<br>2495 | gat gat ggt gat<br>Asp Asp Gly Asp | 7530 |
| gac<br>Asp<br>2500 | aat gcc cct ttg ttt<br>Asn Ala Pro Leu Phe<br>2505 | tac caa cct ggg aaa<br>Tyr Gln Pro Gly Lys<br>2510 | aga gga ttt tat<br>Arg Gly Phe Tyr | 7575 |
| act<br>Thr<br>2515 | cca agg cct ggc aag<br>Pro Arg Pro Gly Lys<br>2520 | aac aca gaa gca agg<br>Asn Thr Glu Ala Arg<br>2525 | ttg aat tgt ttc<br>Leu Asn Cys Phe | 7620 |
| aga<br>Arg<br>2530 | aac att ggc agg att<br>Asn Ile Gly Arg Ile<br>2535 | ctt gga cta tgt ctg<br>Leu Gly Leu Cys Leu<br>2540 | tta cag aat gaa<br>Leu Gln Asn Glu | 7665 |
| ctc<br>Leu<br>2545 | tgt cct atc aca ttg<br>Cys Pro Ile Thr Leu<br>2550 | aat aga cat gta att<br>Asn Arg His Val Ile<br>2555 | aaa gta ttg ctt<br>Lys Val Leu Leu | 7710 |

|      |     |     |     |     |      |     |       |     |     |      |     |     |     |     |  |      |
|------|-----|-----|-----|-----|------|-----|-------|-----|-----|------|-----|-----|-----|-----|--|------|
| ggt  | aga | aaa | gtc | aat | tgg  | cat | gat   | ttt | gct | ttt  | ttt | gat | cct | gta |  | 7755 |
| Gly  | Arg | Lys | Val | Asn | Trp  | His | Asp   | Phe | Ala | Phe  | Phe | Asp | Pro | Val |  |      |
| 2560 |     |     |     |     | 2565 |     |       |     |     | 2570 |     |     |     |     |  |      |
| atg  | tat | gag | agt | ttg | cgg  | caa | cta   | atc | ctc | gcg  | tct | cag | agt | tca |  | 7800 |
| Met  | Tyr | Glu | Ser | Leu | Arg  | Gln | Leu   | Ile | Leu | Ala  | Ser | Gln | Ser | Ser |  |      |
| 2575 |     |     |     |     | 2580 |     |       |     |     | 2585 |     |     |     |     |  |      |
| gat  | gct | gat | gct | gtt | ttc  | tca | gca   | atg | gat | ttg  | gca | ttt | gca | att |  | 7845 |
| Asp  | Ala | Asp | Ala | Val | Phe  | Ser | Ala   | Met | Asp | Leu  | Ala | Phe | Ala | Ile |  |      |
| 2590 |     |     |     |     | 2595 |     |       |     |     | 2600 |     |     |     |     |  |      |
| gac  | ctg | tgt | aaa | gaa | gaa  | ggg | gga   | gga | cag | gtt  | gaa | ctc | att | cct |  | 7890 |
| Asp  | Leu | Cys | Lys | Glu | Glu  | Gly | Gly   | Gly | Gln | Val  | Glu | Leu | Ile | Pro |  |      |
| 2605 |     |     |     |     | 2610 |     |       |     |     | 2615 |     |     |     |     |  |      |
| aat  | ggg | gta | aat | ata | cca  | gtc | act   | cca | cag | aat  | gta | tat | gag | tat |  | 7935 |
| Asn  | Gly | Val | Asn | Ile | Pro  | Val | Thr   | Pro | Gln | Asn  | Val | Tyr | Glu | Tyr |  |      |
| 2620 |     |     |     |     | 2625 |     |       |     |     | 2630 |     |     |     |     |  |      |
| gtg  | cgg | aaa | tac | gca | gaa  | cac | aga   | atg | ttg | gta  | gtt | gca | gaa | cag |  | 7980 |
| Val  | Arg | Lys | Tyr | Ala | Glu  | His | Arg   | Met | Leu | Val  | Val | Ala | Glu | Gln |  |      |
| 2635 |     |     |     |     | 2640 |     |       |     |     | 2645 |     |     |     |     |  |      |
| ccc  | tta | cat | gca | atg | agg  | aaa | ggg   | cta | cta | gat  | gtg | ctt | cca | aaa |  | 8025 |
| Pro  | Leu | His | Ala | Met | Arg  | Lys | Gly   | Leu | Leu | Asp  | Val | Leu | Pro | Lys |  |      |
| 2650 |     |     |     |     | 2655 |     |       |     |     | 2660 |     |     |     |     |  |      |
| aat  | tca | tta | gaa | gat | tta  | acg | gca   | gaa | gat | ttt  | agg | ctt | ttg | gta |  | 8070 |
| Asn  | Ser | Leu | Glu | Asp | Leu  | Thr | Ala   | Glu | Asp | Phe  | Arg | Leu | Leu | Val |  |      |
| 2665 |     |     |     |     | 2670 |     |       |     |     | 2675 |     |     |     |     |  |      |
| aat  | ggc | tgc | ggg | gaa | gtc  | aat | gtg   | caa | atg | ctg  | atc | agt | ttt | acc |  | 8115 |
| Asn  | Gly | Cys | Gly | Glu | Val  | Asn | Val   | Gln | Met | Leu  | Ile | Ser | Phe | Thr |  |      |
| 2680 |     |     |     |     | 2685 |     |       |     |     | 2690 |     |     |     |     |  |      |
| tct  | ttc | aat | gat | gaa | tca  | gga | gaa   | aat | gct | gag  | aag | ctt | ctg | cag |  | 8160 |
| Ser  | Phe | Asn | Asp | Glu | Ser  | Gly | Glu   | Asn | Ala | Glu  | Lys | Leu | Leu | Gln |  |      |
| 2695 |     |     |     |     | 2700 |     |       |     |     | 2705 |     |     |     |     |  |      |
| ttc  | aag | cgt | tgg | ttc | tgg  | tca | ata   | gta | gag | aag  | atg | agc | atg | aca |  | 8205 |
| Phe  | Lys | Arg | Trp | Phe | Trp  | Ser | Ile   | Val | Glu | Lys  | Met | Ser | Met | Thr |  |      |
| 2710 |     |     |     |     | 2715 |     |       |     |     | 2720 |     |     |     |     |  |      |
| gaa  | cga | caa | gat | ctt | gtt  | tac | ttt   | tgg | aca | tca  | agc | cca | tca | ctg |  | 8250 |
| Glu  | Arg | Gln | Asp | Leu | Val  | Tyr | Phe   | Trp | Thr | Ser  | Ser | Pro | Ser | Leu |  |      |
| 2725 |     |     |     |     | 2730 |     |       |     |     | 2735 |     |     |     |     |  |      |
| cca  | gcc | agt | gaa | gaa | gga  | ttc | cag   | cct | atg | ccc  | tca | atc | aca | ata |  | 8295 |
| Pro  | Ala | Ser | Glu | Glu | Gly  | Phe | Gln   | Pro | Met | Pro  | Ser | Ile | Thr | Ile |  |      |
| 2740 |     |     |     |     | 2745 |     |       |     |     | 2750 |     |     |     |     |  |      |
| aga  | cca | cca | gat | gac | caa  | cat | ctt</ |     |     |      |     |     |     |     |  |      |



accttaaaga gataaaatgc agacattcct tgctgagttt atagcttaaa ggctaagga 8610  
gcactagcaa catttggtta tattggtttg ctatgcacca acttctgggt ctaaccccag 8670  
ccaaagatga cagcagaaca acataattta cactgtgatt tatctttttg ctgaggggaa 8730  
aaaaatgtaa atgttctgaa aattcactgc tgcctttgtg gaaactgttt cagcaaaggt 8790  
tcttgtagag agggaatagg gaatttcaaa ataaaaaatt aagtatgt 8838

<210> 64

<211> 2799

<212> PRT

<213> NM\_015902 EDD, hyperplastic discs protein, DD5

<400> 64

Met Thr Ser Ile His Phe Val Val His Pro Leu Pro Gly Thr Glu Asp  
1 5 10 15

Gln Leu Asn Asp Arg Leu Arg Glu Val Ser Glu Lys Leu Asn Lys Tyr  
20 25 30

Asn Leu Asn Ser His Pro Pro Leu Asn Val Leu Glu Gln Ala Thr Ile  
35 40 45

Lys Gln Cys Val Val Gly Pro Asn His Ala Ala Phe Leu Leu Glu Asp  
50 55 60

Gly Arg Val Cys Arg Ile Gly Phe Ser Val Gln Pro Asp Arg Leu Glu  
65 70 75 80

Leu Gly Lys Pro Asp Asn Asn Asp Gly Ser Lys Leu Asn Ser Asn Ser  
85 90 95

Gly Ala Gly Arg Thr Ser Arg Pro Gly Arg Thr Ser Asp Ser Pro Trp  
100 105 110

Phe Leu Ser Gly Ser Glu Thr Leu Gly Arg Leu Ala Gly Asn Thr Leu  
115 120 125

Gly Ser Arg Trp Ser Ser Gly Val Gly Gly Ser Gly Gly Gly Ser Ser  
130 135 140

Gly Arg Ser Ser Ala Gly Ala Arg Asp Ser Arg Arg Gln Thr Arg Val  
145 150 155 160

Ile Arg Thr Gly Arg Asp Arg Gly Ser Gly Leu Leu Gly Ser Gln Pro  
165 170 175

Gln Pro Val Ile Pro Ala Ser Val Ile Pro Glu Glu Leu Ile Ser Gln  
180 185 190

Ala Gln Val Val Leu Gln Gly Lys Ser Arg Ser Val Ile Ile Arg Glu  
195 200 205

Leu Gln Arg Thr Asn Leu Asp Val Asn Leu Ala Val Asn Asn Leu Leu  
210 215 220

Ser Arg Asp Asp Glu Asp Gly Asp Asp Gly Asp Asp Thr Ala Ser Glu  
225 230 235 240

Ser Tyr Leu Pro Gly Glu Asp Leu Met Ser Leu Leu Asp Ala Asp Ile  
245 250 255

His Ser Ala His Pro Ser Val Ile Ile Asp Ala Asp Ala Met Phe Ser  
260 265 270

Glu Asp Ile Ser Tyr Phe Gly Tyr Pro Ser Phe Arg Arg Ser Ser Leu  
275 280 285

Ser Arg Leu Gly Ser Ser Arg Val Leu Leu Leu Pro Leu Glu Arg Asp  
290 295 300

Ser Glu Leu Leu Arg Glu Arg Glu Ser Val Leu Arg Leu Arg Glu Arg  
305 310 315 320

Arg Trp Leu Asp Gly Ala Ser Phe Asp Asn Glu Arg Gly Ser Thr Ser  
325 330 335

Lys Glu Gly Glu Pro Asn Leu Asp Lys Lys Asn Thr Pro Val Gln Ser  
340 345 350

Pro Val Ser Leu Gly Glu Asp Leu Gln Trp Trp Pro Asp Lys Asp Gly  
355 360 365

Thr Lys Phe Ile Cys Ile Gly Ala Leu Tyr Ser Glu Leu Leu Ala Val  
370 375 380

Ser Ser Lys Gly Glu Leu Tyr Gln Trp Lys Trp Ser Glu Ser Glu Pro  
385 390 395 400

Tyr Arg Asn Ala Gln Asn Pro Ser Leu His His Pro Arg Ala Thr Phe  
405 410 415

Leu Gly Leu Thr Asn Glu Lys Ile Val Leu Leu Ser Ala Asn Ser Ile  
420 425 430

Arg Ala Thr Val Ala Thr Glu Asn Asn Lys Val Ala Thr Trp Val Asp  
435 440 445

Glu Thr Leu Ser Ser Val Ala Ser Lys Leu Glu His Thr Ala Gln Thr  
 450 455 460

Tyr Ser Glu Leu Gln Gly Glu Arg Ile Val Ser Leu His Cys Cys Ala  
 465 470 475 480

Leu Tyr Thr Cys Ala Gln Leu Glu Asn Ser Leu Tyr Trp Trp Gly Val  
 485 490 495

Val Pro Phe Ser Gln Arg Lys Lys Met Leu Glu Lys Ala Arg Ala Lys  
 500 505 510

Asn Lys Lys Pro Lys Ser Ser Ala Gly Ile Ser Ser Met Pro Asn Ile  
 515 520 525

Thr Val Gly Thr Gln Val Cys Leu Arg Asn Asn Pro Leu Tyr His Ala  
 530 535 540

Gly Ala Val Ala Phe Ser Ile Ser Ala Gly Ile Pro Lys Val Gly Val  
 545 550 555 560

Leu Met Glu Ser Val Trp Asn Met Asn Asp Ser Cys Arg Phe Gln Leu  
 565 570 575

Arg Ser Pro Glu Ser Leu Lys Asn Met Glu Lys Ala Ser Lys Thr Thr  
 580 585 590

Glu Ala Lys Pro Glu Ser Lys Gln Glu Pro Val Lys Thr Glu Met Gly  
 595 600 605

Pro Pro Pro Ser Pro Ala Ser Thr Cys Ser Asp Ala Ser Ser Ile Ala  
 610 615 620

Ser Ser Ala Ser Met Pro Tyr Lys Arg Arg Arg Ser Thr Pro Ala Pro  
 625 630 635 640

Lys Glu Glu Glu Lys Val Asn Glu Glu Gln Trp Ser Leu Arg Glu Val  
 645 650 655

Val Phe Val Glu Asp Val Lys Asn Val Pro Val Gly Lys Val Leu Lys  
 660 665 670

Val Asp Gly Ala Tyr Val Ala Val Lys Phe Pro Gly Thr Ser Ser Asn  
 675 680 685

Thr Asn Cys Gln Asn Ser Ser Gly Pro Asp Ala Asp Pro Ser Ser Leu  
 690 695 700

Leu Gln Asp Cys Arg Leu Leu Arg Ile Asp Glu Leu Gln Val Val Lys  
 705 710 715 720

Thr Gly Gly Thr Pro Lys Val Pro Asp Cys Phe Gln Arg Thr Pro Lys  
725 730 735

Lys Leu Cys Ile Pro Glu Lys Thr Glu Ile Leu Ala Val Asn Val Asp  
740 745 750

Ser Lys Gly Val His Ala Val Leu Lys Thr Gly Asn Trp Val Arg Tyr  
755 760 765

Cys Ile Phe Asp Leu Ala Thr Gly Lys Ala Glu Gln Glu Asn Asn Phe  
770 775 780

Pro Thr Ser Ser Ile Ala Phe Leu Gly Gln Asn Glu Arg Asn Val Ala  
785 790 795 800

Ile Phe Thr Ala Gly Gln Glu Ser Pro Ile Ile Leu Arg Asp Gly Asn  
805 810 815

Gly Thr Ile Tyr Pro Met Ala Lys Asp Cys Met Gly Gly Ile Arg Asp  
820 825 830

Pro Asp Trp Leu Asp Leu Pro Pro Ile Ser Ser Leu Gly Met Gly Val  
835 840 845

His Ser Leu Ile Asn Leu Pro Ala Asn Ser Thr Ile Lys Lys Lys Ala  
850 855 860

Ala Val Ile Ile Met Ala Val Glu Lys Gln Thr Leu Met Gln His Ile  
865 870 875 880

Leu Arg Cys Asp Tyr Glu Ala Cys Arg Gln Tyr Leu Met Asn Leu Glu  
885 890 895

Gln Ala Val Val Leu Glu Gln Asn Leu Gln Met Leu Gln Thr Phe Ile  
900 905 910

Ser His Arg Cys Asp Gly Asn Arg Asn Ile Leu His Ala Cys Val Ser  
915 920 925

Val Cys Phe Pro Thr Ser Asn Lys Glu Thr Lys Glu Glu Glu Ala  
930 935 940

Glu Arg Ser Glu Arg Asn Thr Phe Ala Glu Arg Leu Ser Ala Val Glu  
945 950 955 960

Ala Ile Ala Asn Ala Ile Ser Val Val Ser Ser Asn Gly Pro Gly Asn  
965 970 975

Arg Ala Gly Ser Ser Ser Arg Ser Leu Arg Leu Arg Glu Met Met  
980 985 990

Arg Arg Ser Leu Arg Ala Ala Gly Leu Gly Arg His Glu Ala Gly Ala  
995 1000 1005

Ser Ser Ser Asp His Gln Asp Pro Val Ser Pro Pro Ile Ala Pro  
1010 1015 1020

Pro Ser Trp Val Pro Asp Pro Pro Ala Met Asp Pro Asp Gly Asp  
1025 1030 1035

Ile Asp Phe Ile Leu Ala Pro Ala Val Gly Ser Leu Thr Thr Ala  
1040 1045 1050

Ala Thr Gly Thr Gly Gln Gly Pro Ser Thr Ser Thr Ile Pro Gly  
1055 1060 1065

Pro Ser Thr Glu Pro Ser Val Val Glu Ser Lys Asp Arg Lys Ala  
1070 1075 1080

Asn Ala His Phe Ile Leu Lys Leu Leu Cys Asp Ser Val Val Leu  
1085 1090 1095

Gln Pro Tyr Leu Arg Glu Leu Leu Ser Ala Lys Asp Ala Arg Gly  
1100 1105 1110

Met Thr Pro Phe Met Ser Ala Val Ser Gly Arg Ala Tyr Pro Ala  
1115 1120 1125

Ala Ile Thr Ile Leu Glu Thr Ala Gln Lys Ile Ala Lys Ala Glu  
1130 1135 1140

Ile Ser Ser Ser Glu Lys Glu Glu Asp Val Phe Met Gly Met Val  
1145 1150 1155

Cys Pro Ser Gly Thr Asn Pro Asp Asp Ser Pro Leu Tyr Val Leu  
1160 1165 1170

Cys Cys Asn Asp Thr Cys Ser Phe Thr Trp Thr Gly Ala Glu His  
1175 1180 1185

Ile Asn Gln Asp Ile Phe Glu Cys Arg Thr Cys Gly Leu Leu Glu  
1190 1195 1200

Ser Leu Cys Cys Cys Thr Glu Cys Ala Arg Val Cys His Lys Gly  
1205 1210 1215

His Asp Cys Lys Leu Lys Arg Thr Ser Pro Thr Ala Tyr Cys Asp  
1220 1225 1230

Cys Trp Glu Lys Cys Lys Cys Lys Thr Leu Ile Ala Gly Gln Lys

| 1235            | 1240                        | 1245                                    |
|-----------------|-----------------------------|---|
| Ser Ala<br>1250 | Arg Leu Asp Leu Leu<br>1255 | Tyr Arg Leu Leu Thr Ala Thr Asn<br>1260 |
| Leu Val<br>1265 | Thr Leu Pro Asn Ser<br>1270 | Arg Gly Glu His Leu Leu Leu Phe<br>1275 |
| Leu Val<br>1280 | Gln Thr Val Ala Arg<br>1285 | Gln Thr Val Glu His Cys Gln Tyr<br>1290 |
| Arg Pro<br>1295 | Pro Arg Ile Arg Glu<br>1300 | Asp Arg Asn Arg Lys Thr Ala Ser<br>1305 |
| Pro Glu<br>1310 | Asp Ser Asp Met Pro<br>1315 | Asp His Asp Leu Glu Pro Pro Arg<br>1320 |
| Phe Ala<br>1325 | Gln Leu Ala Leu Glu<br>1330 | Arg Val Leu Gln Asp Trp Asn Ala<br>1335 |
| Leu Lys<br>1340 | Ser Met Ile Met Phe<br>1345 | Gly Ser Gln Glu Asn Lys Asp Pro<br>1350 |
| Leu Ser<br>1355 | Ala Ser Ser Arg Ile<br>1360 | Gly His Leu Leu Pro Glu Glu Gln<br>1365 |
| Val Tyr<br>1370 | Leu Asn Gln Gln Ser<br>1375 | Gly Thr Ile Arg Leu Asp Cys Phe<br>1380 |
| Thr His<br>1385 | Cys Leu Ile Val Lys<br>1390 | Cys Thr Ala Asp Ile Leu Leu Leu<br>1395 |
| Asp Thr<br>1400 | Leu Leu Gly Thr Leu<br>1405 | Val Lys Glu Leu Gln Asn Lys Tyr<br>1410 |
| Thr Pro<br>1415 | Gly Arg Arg Glu Glu<br>1420 | Ala Ile Ala Val Thr Met Arg Phe<br>1425 |
| Leu Arg<br>1430 | Ser Val Ala Arg Val<br>1435 | Phe Val Ile Leu Ser Val Glu Met<br>1440 |
| Ala Ser<br>1445 | Ser Lys Lys Lys Asn<br>1450 | Asn Phe Ile Pro Gln Pro Ile Gly<br>1455 |
| Lys Cys<br>1460 | Lys Arg Val Phe Gln<br>1465 | Ala Leu Leu Pro Tyr Ala Val Glu<br>1470 |
| Glu Leu<br>1475 | Cys Asn Val Ala Glu<br>1480 | Ser Leu Ile Val Pro Val Arg Met<br>1485 |

Gly Ile Ala Arg Pro Thr Ala Pro Phe Thr Leu Ala Ser Thr Ser  
1490 1495 1500

Ile Asp Ala Met Gln Gly Ser Glu Glu Leu Phe Ser Val Glu Pro  
1505 1510 1515

Leu Pro Pro Arg Pro Ser Ser Asp Gln Ser Ser Ser Ser Ser Gln  
1520 1525 1530

Ser Gln Ser Ser Tyr Ile Ile Arg Asn Pro Gln Gln Arg Arg Ile  
1535 1540 1545

Ser Gln Ser Gln Pro Val Arg Gly Arg Asp Glu Glu Gln Asp Asp  
1550 1555 1560

Ile Val Ser Ala Asp Val Glu Glu Val Glu Val Val Glu Gly Val  
1565 1570 1575

Ala Gly Glu Glu Asp His His Asp Glu Gln Glu Glu His Gly Glu  
1580 1585 1590

Glu Asn Ala Glu Ala Glu Gly Gln His Asp Glu His Asp Glu Asp  
1595 1600 1605

Gly Ser Asp Met Glu Leu Asp Leu Leu Ala Ala Ala Glu Thr Glu  
1610 1615 1620

Ser Asp Ser Glu Ser Asn His Ser Asn Gln Asp Asn Ala Ser Gly  
1625 1630 1635

Arg Arg Ser Val Val Thr Ala Ala Thr Ala Gly Ser Glu Ala Gly  
1640 1645 1650

Ala Ser Ser Val Pro Ala Phe Phe Ser Glu Asp Asp Ser Gln Ser  
1655 1660 1665

Asn Asp Ser Ser Asp Ser Asp Ser Ser Ser Ser Gln Ser Asp Asp  
1670 1675 1680

Ile Glu Gln Glu Thr Phe Met Leu Asp Glu Pro Leu Glu Arg Thr  
1685 1690 1695

Thr Asn Ser Ser His Ala Asn Gly Ala Ala Gln Ala Pro Arg Ser  
1700 1705 1710

Met Gln Trp Ala Val Arg Asn Thr Gln His Gln Arg Ala Ala Ser  
1715 1720 1725

Thr Ala Pro Ser Ser Thr Ser Thr Pro Ala Ala Ser Ser Ala Gly  
1730 1735 1740

Leu Ile Tyr Ile Asp Pro Ser Asn Leu Arg Arg Ser Gly Thr Ile  
1745 1750 1755

Ser Thr Ser Ala Ala Ala Ala Ala Ala Ala Leu Glu Ala Ser Asn  
1760 1765 1770

Ala Ser Ser Tyr Leu Thr Ser Ala Ser Ser Leu Ala Arg Ala Tyr  
1775 1780 1785

Ser Ile Val Ile Arg Gln Ile Ser Asp Leu Met Gly Leu Ile Pro  
1790 1795 1800

Lys Tyr Asn His Leu Val Tyr Ser Gln Ile Pro Ala Ala Val Lys  
1805 1810 1815

Leu Thr Tyr Gln Asp Ala Val Asn Leu Gln Asn Tyr Val Glu Glu  
1820 1825 1830

Lys Leu Ile Pro Thr Trp Asn Trp Met Val Ser Ile Met Asp Ser  
1835 1840 1845

Thr Glu Ala Gln Leu Arg Tyr Gly Ser Ala Leu Ala Ser Ala Gly  
1850 1855 1860

Asp Pro Gly His Pro Asn His Pro Leu His Ala Ser Gln Asn Ser  
1865 1870 1875

Ala Arg Arg Glu Arg Met Thr Ala Arg Glu Glu Ala Ser Leu Arg  
1880 1885 1890

Thr Leu Glu Gly Arg Arg Arg Ala Thr Leu Leu Ser Ala Arg Gln  
1895 1900 1905

Gly Met Met Ser Ala Arg Gly Asp Phe Leu Asn Tyr Ala Leu Ser  
1910 1915 1920

Leu Met Arg Ser His Asn Asp Glu His Ser Asp Val Leu Pro Val  
1925 1930 1935

Leu Asp Val Cys Ser Leu Lys His Val Ala Tyr Val Phe Gln Ala  
1940 1945 1950

Leu Ile Tyr Trp Ile Lys Ala Met Asn Gln Gln Thr Thr Leu Asp  
1955 1960 1965

Thr Pro Gln Leu Glu Arg Lys Arg Thr Arg Glu Leu Leu Glu Leu  
1970 1975 1980

Gly Ile Asp Asn Glu Asp Ser Glu His Glu Asn Asp Asp Asp Thr  
1985 1990 1995



Asn Gln Ser Ala Thr Leu Asn Asp Lys Asp Asp Asp Ser Leu Pro  
2000 2005 2010

Ala Glu Thr Gly Gln Asn His Pro Phe Phe Arg Arg Ser Asp Ser  
2015 2020 2025

Met Thr Phe Leu Gly Cys Ile Pro Pro Asn Pro Phe Glu Val Pro  
2030 2035 2040

Leu Ala Glu Ala Ile Pro Leu Ala Asp Gln Pro His Leu Leu Gln  
2045 2050 2055

Pro Asn Ala Arg Lys Glu Asp Leu Phe Gly Arg Pro Ser Gln Gly  
2060 2065 2070

Leu Tyr Ser Ser Ser Ala Ser Ser Gly Lys Cys Leu Met Glu Val  
2075 2080 2085

Thr Val Asp Arg Asn Cys Leu Glu Val Leu Pro Thr Lys Met Ser  
2090 2095 2100

Tyr Ala Ala Asn Leu Lys Asn Val Met Asn Met Gln Asn Arg Gln  
2105 2110 2115

Lys Lys Glu Gly Glu Glu Gln Pro Val Leu Pro Glu Glu Thr Glu  
2120 2125 2130

Ser Ser Lys Pro Gly Pro Ser Ala His Asp Leu Ala Ala Gln Leu  
2135 2140 2145

Lys Ser Ser Leu Leu Ala Glu Ile Gly Leu Thr Glu Ser Glu Gly  
2150 2155 2160

Pro Pro Leu Thr Ser Phe Arg Pro Gln Cys Ser Phe Met Gly Met  
2165 2170 2175

Val Ile Ser His Asp Met Leu Leu Gly Arg Trp Arg Leu Ser Leu  
2180 2185 2190

Glu Leu Phe Gly Arg Val Phe Met Glu Asp Val Gly Ala Glu Pro  
2195 2200 2205

Gly Ser Ile Leu Thr Glu Leu Gly Gly Phe Glu Val Lys Glu Ser  
2210 2215 2220

Lys Phe Arg Arg Glu Met Glu Lys Leu Arg Asn Gln Gln Ser Arg  
2225 2230 2235

Asp Leu Ser Leu Glu Val Asp Arg Asp Arg Asp Leu Leu Ile Gln  
2240 2245 2250

Gln Thr Met Arg Gln Leu Asn Asn His Phe Gly Arg Arg Cys Ala  
2255 2260 2265

Thr Thr Pro Met Ala Val His Arg Val Lys Val Thr Phe Lys Asp  
2270 2275 2280

Glu Pro Gly Glu Gly Ser Gly Val Ala Arg Ser Phe Tyr Thr Ala  
2285 2290 2295

Ile Ala Gln Ala Phe Leu Ser Asn Glu Lys Leu Pro Asn Leu Glu  
2300 2305 2310

Cys Ile Gln Asn Ala Asn Lys Gly Thr His Thr Ser Leu Met Gln  
2315 2320 2325

Arg Leu Arg Asn Arg Gly Glu Arg Asp Arg Glu Arg Glu Arg Glu  
2330 2335 2340

Arg Glu Met Arg Arg Ser Ser Gly Leu Arg Ala Gly Ser Arg Arg  
2345 2350 2355

Asp Arg Asp Arg Asp Phe Arg Arg Gln Leu Ser Ile Asp Thr Arg  
2360 2365 2370

Pro Phe Arg Pro Ala Ser Glu Gly Asn Pro Ser Asp Asp Pro Glu  
2375 2380 2385

Pro Leu Pro Ala His Arg Gln Ala Leu Gly Glu Arg Leu Tyr Pro  
2390 2395 2400

Arg Val Gln Ala Met Gln Pro Ala Phe Ala Ser Lys Ile Thr Gly  
2405 2410 2415

Met Leu Leu Glu Leu Ser Pro Ala Gln Leu Leu Leu Leu Leu Ala  
2420 2425 2430

Ser Glu Asp Ser Leu Arg Ala Arg Val Asp Glu Ala Met Glu Leu  
2435 2440 2445

Ile Ile Ala His Gly Arg Glu Asn Gly Ala Asp Ser Ile Leu Asp  
2450 2455 2460

Leu Gly Leu Val Asp Ser Ser Glu Lys Val Gln Gln Glu Asn Arg  
2465 2470 2475

Lys Arg His Gly Ser Ser Arg Ser Val Val Asp Met Asp Leu Asp  
2480 2485 2490

Asp Thr Asp Asp Gly Asp Asp Asn Ala Pro Leu Phe Tyr Gln Pro  
2495 2500 2505

Gly Lys Arg Gly Phe Tyr Thr Pro Arg Pro Gly Lys Asn Thr Glu  
2510 2515 2520

Ala Arg Leu Asn Cys Phe Arg Asn Ile Gly Arg Ile Leu Gly Leu  
2525 2530 2535

Cys Leu Leu Gln Asn Glu Leu Cys Pro Ile Thr Leu Asn Arg His  
2540 2545 2550

Val Ile Lys Val Leu Leu Gly Arg Lys Val Asn Trp His Asp Phe  
2555 2560 2565

Ala Phe Phe Asp Pro Val Met Tyr Glu Ser Leu Arg Gln Leu Ile  
2570 2575 2580

Leu Ala Ser Gln Ser Ser Asp Ala Asp Ala Val Phe Ser Ala Met  
2585 2590 2595

Asp Leu Ala Phe Ala Ile Asp Leu Cys Lys Glu Glu Gly Gly Gly  
2600 2605 2610

Gln Val Glu Leu Ile Pro Asn Gly Val Asn Ile Pro Val Thr Pro  
2615 2620 2625

Gln Asn Val Tyr Glu Tyr Val Arg Lys Tyr Ala Glu His Arg Met  
2630 2635 2640

Leu Val Val Ala Glu Gln Pro Leu His Ala Met Arg Lys Gly Leu  
2645 2650 2655

Leu Asp Val Leu Pro Lys Asn Ser Leu Glu Asp Leu Thr Ala Glu  
2660 2665 2670

Asp Phe Arg Leu Leu Val Asn Gly Cys Gly Glu Val Asn Val Gln  
2675 2680 2685

Met Leu Ile Ser Phe Thr Ser Phe Asn Asp Glu Ser Gly Glu Asn  
2690 2695 2700

Ala Glu Lys Leu Leu Gln Phe Lys Arg Trp Phe Trp Ser Ile Val  
2705 2710 2715

Glu Lys Met Ser Met Thr Glu Arg Gln Asp Leu Val Tyr Phe Trp  
2720 2725 2730

Thr Ser Ser Pro Ser Leu Pro Ala Ser Glu Glu Gly Phe Gln Pro  
2735 2740 2745

Met Pro Ser Ile Thr Ile Arg Pro Pro Asp Asp Gln His Leu Pro  
2750 2755 2760

Thr Ala Asn Thr Cys Ile Sér Arg Leu Tyr Val Pro Leu Tyr Sér  
2765 2770 2775

Ser Lys Gln Ile Leu Lys Gln Lys Leu Leu Leu Ala Ile Lys Thr  
2780 2785 2790

Lys Asn Phe Gly Phe Val  
2795

<210> 65

<211> 4256

<212> DNA

<213> NM\_014992 DAAM1

<220>

<221> CDS

<222> (126)..(3359)

<223>

<400> 65

tacgagcggc tgtgaaggaa actgtttaac cggatcccat tgtaccaga gtgcagagcc 60

gcctttccag catgcagggg ctgctcagcg ttttagtcaca tcaagaaata gaacagaatt 120

cagcc atg gcc cca aga aag aga ggt gga cga ggt att tca ttc atc ttt 170  
Met Ala Pro Arg Lys Arg Gly Gly Arg Ile Ser Phe Ile Phe  
1 5 10 15

tgc tgt ttc cga aat aat gat cac cca gaa atc acg tat cgg ctg cga 218  
Cys Cys Phe Arg Asn Asn Asp His Pro Glu Ile Thr Tyr Arg Leu Arg  
20 25 30

aat gat agc aac ttt gcg ctt cag acc atg gaa cca gca ttg ccc atg 266  
Asn Asp Ser Asn Phe Ala Leu Gln Thr Met Glu Pro Ala Leu Pro Met  
35 40 45

ccc cct gtg gag gag ctg gat gtc atg ttc agt gaa ctg gtg gat gaa 314  
Pro Pro Val Glu Glu Leu Asp Val Met Phe Ser Glu Leu Val Asp Glu  
50 55 60

ctg gac ctc aca gac aaa cac aga gaa gcc atg ttt gca ctt cca gct 362  
Leu Asp Leu Thr Asp Lys His Arg Glu Ala Met Phe Ala Leu Pro Ala  
65 70 75

gag aaa aaa tgg caa ata tac tgt agc aag aaa aag gac cag gaa gaa 410  
Glu Lys Lys Trp Gln Ile Tyr Cys Ser Lys Lys Lys Asp Gln Glu Glu  
80 85 90 95

aac aag gga gct aca agt tgg cct gaa ttc tac att gat cag ctc aat 458  
Asn Lys Gly Ala Thr Ser Trp Pro Glu Phe Tyr Ile Asp Gln Leu Asn  
100 105 110

tcc atg gct gct aga aaa tct ctg ctg gct tta gag aag gaa gaa gaa 506  
Ser Met Ala Ala Arg Lys Ser Leu Leu Ala Leu Glu Lys Glu Glu Glu

| 115   | 120 | 125 |      |
|---|-----|-----|------|
| gaa gaa aga agt aaa act ata gag agt tta aag aca gca ctg agg aca<br>Glu Glu Arg Ser Lys Thr Ile Glu Ser Leu Lys Thr Ala Leu Arg Thr<br>130 135 140     |     |     | 554  |
| aaa cca atg agg ttt gta acc aga ttc atc gac ttg gat ggc cta tca<br>Lys Pro Met Arg Phe Val Thr Arg Phe Ile Asp Leu Asp Gly Leu Ser<br>145 150 155     |     |     | 602  |
| tgt atc ctc aac ttt cta aag acc atg gac tac gag acc tca gag tct<br>Cys Ile Leu Asn Phe Leu Lys Thr Met Asp Tyr Glu Thr Ser Glu Ser<br>160 165 170 175 |     |     | 650  |
| cga ata cat act tct ctc att ggc tgt ata aag gcg tta atg aac aac<br>Arg Ile His Thr Ser Leu Ile Gly Cys Ile Lys Ala Leu Met Asn Asn<br>180 185 190     |     |     | 698  |
| tct caa ggc cgg gct cac gtc ctg gct cat tct gag agt att aat gta<br>Ser Gln Gly Arg Ala His Val Leu Ala His Ser Glu Ser Ile Asn Val<br>195 200 205     |     |     | 746  |
| att gct cag agt ctg agc aca gag aac att aaa acg aag gtg gcc gtg<br>Ile Ala Gln Ser Leu Ser Thr Glu Asn Ile Lys Thr Lys Val Ala Val<br>210 215 220     |     |     | 794  |
| ctg gaa atc ttg ggc gcc gtg tgc ctg gtt ccc ggg ggc cac aag aag<br>Leu Glu Ile Leu Gly Ala Val Cys Leu Val Pro Gly Gly His Lys Lys<br>225 230 235     |     |     | 842  |
| gtt ctg cag gcc atg ctg cac tac cag aag tat gcc agc gaa agg acc<br>Val Leu Gln Ala Met Leu His Tyr Gln Lys Tyr Ala Ser Glu Arg Thr<br>240 245 250 255 |     |     | 890  |
| cgc ttt cag aca tta att aac gac ttg gat aaa agc act ggg cgg tat<br>Arg Phe Gln Thr Leu Ile Asn Asp Leu Asp Lys Ser Thr Gly Arg Tyr<br>260 265 270     |     |     | 938  |
| cga gat gaa gtg agt ctc aag act gcc atc atg tcc ttc att aat gca<br>Arg Asp Glu Val Ser Leu Lys Thr Ala Ile Met Ser Phe Ile Asn Ala<br>275 280 285     |     |     | 986  |
| gtg ctc agc caa ggt gca gga gtg gag agt ttg gac ttt aga ctt cat<br>Val Leu Ser Gln Gly Ala Gly Val Glu Ser Leu Asp Phe Arg Leu His<br>290 295 300     |     |     | 1034 |
| ctt cgc tat gaa ttt ctg atg tta gga att caa cct gta ata gat aaa<br>Leu Arg Tyr Glu Phe Leu Met Leu Gly Ile Gln Pro Val Ile Asp Lys<br>305 310 315     |     |     | 1082 |
| tta agg gaa cac gaa aat tca aca tta gat agg cat tta gac ttt ttt<br>Leu Arg Glu His Glu Asn Ser Thr Leu Asp Arg His Leu Asp Phe Phe<br>320 325 330 335 |     |     | 1130 |
| gaa atg ctc cga aat gaa gat gaa cta gaa ttt gcc aaa aga ttt gaa<br>Glu Met Leu Arg Asn Glu Asp Glu Leu Glu Phe Ala Lys Arg Phe Glu<br>340 345 350     |     |     | 1178 |
| ctg gtt cac ata gac aca aaa agt gca act cag atg ttt gag ctg acc<br>Leu Val His Ile Asp Thr Lys Ser Ala Thr Gln Met Phe Glu Leu Thr<br>355 360 365     |     |     | 1226 |
| agg aag agg ctg aca cat agt gaa gct tac ccg cat ttc atg tcc atc<br>Arg Lys Arg Leu Thr His Ser Glu Ala Tyr Pro His Phe Met Ser Ile<br>370 375 380     |     |     | 1274 |
| ctg cac cac tgc ctc caa atg cct tac aag agg agt ggc aac act gtt<br>Leu His His Cys Leu Gln Met Pro Tyr Lys Arg Ser Gly Asn Thr Val                    |     |     | 1322 |

| 385   | 390 | 395 |      |
|---|-----|-----|------|
| cag tac tgg cta cta cta gat aga att ata cag cag ata gtt atc cag<br>Gln Tyr Trp Leu Leu Leu Asp Arg Ile Ile Gln Gln Ile Val Ile Gln<br>400 405 410 415 |     |     | 1370 |
| aat gac aaa gga cag gac cct gac tcc aca cct ttg gaa aac ttt aat<br>Asn Asp Lys Gly Gln Asp Pro Asp Ser Thr Pro Leu Glu Asn Phe Asn<br>420 425 430     |     |     | 1418 |
| att aag aat gtc gta cga atg ttg gtt aat gaa aat gaa gtt aag cag<br>Ile Lys Asn Val Val Arg Met Leu Val Asn Glu Asn Glu Val Lys Gln<br>435 440 445     |     |     | 1466 |
| tgg aaa gaa caa gcg gaa aaa atg aga aaa gag cac aat gag cta caa<br>Trp Lys Glu Gln Ala Glu Lys Met Arg Lys Glu His Asn Glu Leu Gln<br>450 455 460     |     |     | 1514 |
| cag aaa ctg gaa aag aaa gaa cga gaa tgt gat gct aag act caa gag<br>Gln Lys Leu Glu Lys Lys Glu Arg Glu Cys Asp Ala Lys Thr Gln Glu<br>465 470 475     |     |     | 1562 |
| aag gaa gag atg atg cag acc tta aat aaa atg aaa gag aaa ctt gaa<br>Lys Glu Glu Met Met Gln Thr Leu Asn Lys Met Lys Glu Lys Leu Glu<br>480 485 490 495 |     |     | 1610 |
| aag gag act act gag cat aag caa gtc aag cag cag gtg gcg gac ctc<br>Lys Glu Thr Thr Glu His Lys Lys Gln Val Lys Gln Val Ala Asp Leu<br>500 505 510     |     |     | 1658 |
| aca gca cag ctc cat gag ctc agc agg agg gcc gtc tgt gct tca atc<br>Thr Ala Gln Leu His Glu Leu Ser Arg Arg Ala Val Cys Ala Ser Ile<br>515 520 525     |     |     | 1706 |
| cca ggt gga ccc tcg cct gga gca cca gga ggg ccc ttt cct tcc tct<br>Pro Gly Gly Pro Ser Pro Gly Ala Pro Gly Gly Pro Phe Pro Ser Ser<br>530 535 540     |     |     | 1754 |
| gtg cct gga tct ctc ctt cct ccc cca cca ccc cca cct cta cca ggt<br>Val Pro Gly Ser Leu Leu Pro Pro Pro Pro Pro Pro Leu Pro Gly<br>545 550 555         |     |     | 1802 |
| ggg atg ctt ccc cct cca ccg cct ccc ctc cct cca ggt ggc cct cct<br>Gly Met Leu Pro Pro Pro Pro Pro Pro Leu Pro Gly Gly Pro Pro<br>560 565 570 575     |     |     | 1850 |
| cct ccc cca ggg cct cct ccc tta ggg gca atc atg cca cct cct ggt<br>Pro Pro Pro Gly Pro Pro Pro Leu Gly Ala Ile Met Pro Pro Pro Gly<br>580 585 590     |     |     | 1898 |
| gct cca atg ggc cta gca ctg aag aag aaa agc att cct cag ccc aca<br>Ala Pro Met Gly Leu Ala Leu Lys Lys Lys Ser Ile Pro Gln Pro Thr<br>595 600 605     |     |     | 1946 |
| aat gcc ctg aaa tcc ttc aac tgg tct aaa ctg ccc gag aac aaa ctg<br>Asn Ala Leu Lys Ser Phe Asn Trp Ser Lys Leu Pro Glu Asn Lys Leu<br>610 615 620     |     |     | 1994 |
| gaa gga aca gta tgg acc gaa att gat gat aca aaa gtc ttc aaa att<br>Glu Gly Thr Val Trp Thr Glu Ile Asp Asp Thr Lys Val Phe Lys Ile<br>625 630 635     |     |     | 2042 |
| cta gat ctt gaa gac ctg gaa aga acc ttc tct gcc tat caa aga cag<br>Leu Asp Leu Glu Asp Leu Glu Arg Thr Phe Ser Ala Tyr Gln Arg Gln<br>640 645 650 655 |     |     | 2090 |
| cag gat ttc ttt gtg aac agt aac tcc aag cag aaa gaa gca gat gcc<br>Gln Asp Phe Phe Val Asn Ser Asn Ser Lys Gln Lys Glu Ala Asp Ala<br>660 665 670 675 |     |     | 2138 |

| 660   | 665 | 670 |      |
|---|-----|-----|------|
| att gat gac act ctg agt tcc aaa ctt aaa gtt aaa gag ctt tcg gtg<br>Ile Asp Asp Thr Leu Ser Ser Lys Leu Lys Val Lys Glu Leu Ser Val<br>675 680 685     |     |     | 2186 |
| att gat ggt cgg aga gct cag aat tgc aac atc ctt cta tcg agg ttg<br>Ile Asp Gly Arg Arg Ala Gln Asn Cys Asn Ile Leu Leu Ser Arg Leu<br>690 695 700     |     |     | 2234 |
| aaa tta tcc aat gac gaa atc aaa cgg gca att cta aca atg gac gaa<br>Lys Leu Ser Asn Asp Glu Ile Lys Arg Ala Ile Leu Thr Met Asp Glu<br>705 710 715     |     |     | 2282 |
| cag gaa gat ctg ccc aag gac atg ttg gaa cag ctc ttg aaa ttt gtt<br>Gln Glu Asp Leu Pro Lys Asp Met Leu Glu Gln Leu Leu Lys Phe Val<br>720 725 730 735 |     |     | 2330 |
| cct gaa aaa agt gac att gac cta ttg gag gaa cat aaa cac gaa ctg<br>Pro Glu Lys Ser Asp Ile Asp Leu Leu Glu Glu His Lys His Glu Leu<br>740 745 750     |     |     | 2378 |
| gat cgg atg gcc aag gct gat agg ttc ctt ttt gag atg agc cga att<br>Asp Arg Met Ala Lys Ala Asp Arg Phe Leu Phe Glu Met Ser Arg Ile<br>755 760 765     |     |     | 2426 |
| aat cac tat cag caa agg ttg caa tcg ctg tac ttc aaa aag aag ttt<br>Asn His Tyr Gln Gln Arg Leu Gln Ser Leu Tyr Phe Lys Lys Lys Phe<br>770 775 780     |     |     | 2474 |
| gca gag cgt gtg gca gaa gtg aaa cct aaa gtg gaa gca att cgt tct<br>Ala Glu Arg Val Ala Glu Val Lys Pro Lys Val Glu Ala Ile Arg Ser<br>785 790 795     |     |     | 2522 |
| ggc tca gaa gag gtg ttt agg agt ggt gcc ctc aag cag ttg ctg gag<br>Gly Ser Glu Glu Val Phe Arg Ser Gly Ala Leu Lys Gln Leu Leu Glu<br>800 805 810 815 |     |     | 2570 |
| gtg gtt ttg gca ttt gga aat tat atg aat aaa ggt caa aga ggg aat<br>Val Val Leu Ala Phe Gly Asn Tyr Met Asn Lys Gly Gln Arg Gly Asn<br>820 825 830     |     |     | 2618 |
| gca tat gga ttc aag ata tct agc cta aac aaa att gct gac aca aaa<br>Ala Tyr Gly Phe Lys Ile Ser Ser Leu Asn Lys Ile Ala Asp Thr Lys<br>835 840 845     |     |     | 2666 |
| tcc agc atc gac aaa aac att acc ctt ttg cac tat ctc atc act att<br>Ser Ser Ile Asp Lys Asn Ile Thr Leu Leu His Tyr Leu Ile Thr Ile<br>850 855 860     |     |     | 2714 |
| gtg gaa aat aag tac ccc agt gtt ctc aat cta aat gaa gaa ttg cga<br>Val Glu Asn Lys Tyr Pro Ser Val Leu Asn Leu Asn Glu Glu Leu Arg<br>865 870 875     |     |     | 2762 |
| gat att cct caa gct gcg aaa gta aac atg act gag ctg gac aaa gaa<br>Asp Ile Pro Gln Ala Ala Lys Val Asn Met Thr Glu Leu Asp Lys Glu<br>880 885 890 895 |     |     | 2810 |
| ata agt acc ttg aga agt ggc ttg aaa gca gta gag aca gag ctg gaa<br>Ile Ser Thr Leu Arg Ser Gly Leu Lys Ala Val Glu Thr Glu Leu Glu<br>900 905 910     |     |     | 2858 |
| tat cag aag tct cag ccc cca cag ccc gga gat aag ttt gtg tct gtt<br>Tyr Gln Lys Ser Gln Pro Pro Gln Pro Gly Asp Lys Phe Val Ser Val<br>915 920 925     |     |     | 2906 |
| gtc agc cag ttc atc aca gta gcc agc ttc agc ttc tct gat gtt gaa<br>Val Ser Gln Phe Ile Thr Val Ala Ser Phe Ser Phe Ser Asp Val Glu<br>930 935 940 945 |     |     | 2954 |

| 930   | 935 | 940 |      |
|---|-----|-----|------|
| gac ctt cta gca gaa gct aaa gac ctg ttt act aaa gca gtg aag cac<br>Asp Leu Leu Ala Glu Ala Lys Asp Leu Phe Thr Lys Ala Val Lys His<br>945 950 955     |     |     | 3002 |
| ttt ggg gaa gag gct ggc aaa ata caa cca gat gag ttc ttt ggc att<br>Phe Gly Glu Glu Ala Gly Lys Ile Gln Pro Asp Glu Phe Phe Gly Ile<br>960 965 970 975 |     |     | 3050 |
| ttt gat caa ttt ctt caa gct gtg tca gaa gcc aaa caa gaa aac gaa<br>Phe Asp Gln Phe Leu Gln Ala Val Ser Glu Ala Lys Gln Glu Asn Glu<br>980 985 990     |     |     | 3098 |
| aat atg aga aag aaa aag gag gaa gaa gaa cgt cga gct cgc atg gaa<br>Asn Met Arg Lys Lys Lys Glu Glu Glu Glu Arg Arg Ala Arg Met Glu<br>995 1000 1005   |     |     | 3146 |
| gct cag ctc aaa gaa caa cgt gaa agg gaa cgt aaa atg aga aaa<br>Ala Gln Leu Lys Glu Gln Arg Glu Arg Glu Arg Lys Met Arg Lys<br>1010 1015 1020          |     |     | 3191 |
| gct aaa gag aat agt gaa gaa agc gga gag ttt gat gac ctt gtt<br>Ala Lys Glu Asn Ser Glu Glu Ser Gly Glu Phe Asp Asp Leu Val<br>1025 1030 1035          |     |     | 3236 |
| tca gct tta cgc tca gga gaa gtg ttt gac aaa gac ctt tct aaa<br>Ser Ala Leu Arg Ser Gly Glu Val Phe Asp Lys Asp Leu Ser Lys<br>1040 1045 1050          |     |     | 3281 |
| ttg aaa cgg aat cgc aaa cgt att acc aac cag atg act gac agc<br>Leu Lys Arg Asn Arg Lys Arg Ile Thr Asn Gln Met Thr Asp Ser<br>1055 1060 1065          |     |     | 3326 |
| agc aga gag aga cca atc aca aaa ctt aat ttc taattttcca<br>Ser Arg Glu Arg Pro Ile Thr Lys Leu Asn Phe<br>1070 1075                                    |     |     | 3369 |
| tgaatacttt ttttttagaaa gctcattagc agccctctaa agtgactaga acgtttcatt  |     |     | 3429 |
| acactgcctt gcaatccaaa cagtggcaat tttttccttc atctgtgagt gaatgtgtga   |     |     | 3489 |
| acgtgtgtat gtaaatgtat gtgtgtatat attaaaaat gtatatagat gtctgagtgt  |     |     | 3549 |
| tgctctggaga cctatacgtat tgggttaaaaa gatttatgtt aatgtatgtg ctccaaaacc  |     |     | 3609 |
| tttcgtgtat gcattcacat tgagtgtggc tcattttctt tccccgaacg ccatgactgt   |     |     | 3669 |
| tcagaagcac aatactatct cctgaaagag ataagagaca ttccctagat tcaaaggcaa   |     |     | 3729 |
| aacagaagaa acaaacaac aaacaacaa agcttgcaaa atattttatg gtttccaagc   |     |     | 3789 |
| ttgatatect ttaaaattat ttccattgat ggaactggag ttgttggaag aacatagatt   |     |     | 3849 |
| taaaatgatt ttgatagct gacattgtga tgttgatgta tcacatcagt aataggacca  |     |     | 3909 |
| gctttgaatt tctgacattg gtgtggggat acagtctgta aatgtttatt gagaacatct   |     |     | 3969 |
| tgacacacaa ttgaattatg tagaatgtca atcaagtttt tgtatattta aaagttggac   |     |     | 4029 |
| atcaattttt tcccctgatt tcatcaagtt atctctgcca agtgcctctg ataatttctt   |     |     | 4089 |
| cagatttttg gaaaaaaca ctatataaat gcaatccatg ctttttttaa agaacaacat  |     |     | 4149 |
| tgccagagta tgcttgttct aacaatatag atatataaac cttaaaaata ataaaatctc   |     |     | 4209 |
| tcaccaaga cttaaaggaa gaattctctg aagggataaa gattact  |     |     | 4256 |



&lt;210&gt; 66

&lt;211&gt; 1078

&lt;212&gt; PRT

&lt;213&gt; NM\_014992 DAAM1

&lt;400&gt; 66

Met Ala Pro Arg Lys Arg Gly Gly Arg Gly Ile Ser Phe Ile Phe Cys  
1 5 10 15

Cys Phe Arg Asn Asn Asp His Pro Glu Ile Thr Tyr Arg Leu Arg Asn  
20 25 30

Asp Ser Asn Phe Ala Leu Gln Thr Met Glu Pro Ala Leu Pro Met Pro  
35 40 45

Pro Val Glu Glu Leu Asp Val Met Phe Ser Glu Leu Val Asp Glu Leu  
50 55 60

Asp Leu Thr Asp Lys His Arg Glu Ala Met Phe Ala Leu Pro Ala Glu  
65 70 75 80

Lys Lys Trp Gln Ile Tyr Cys Ser Lys Lys Lys Asp Gln Glu Glu Asn  
85 90 95

Lys Gly Ala Thr Ser Trp Pro Glu Phe Tyr Ile Asp Gln Leu Asn Ser  
100 105 110

Met Ala Ala Arg Lys Ser Leu Leu Ala Leu Glu Lys Glu Glu Glu Glu  
115 120 125

Glu Arg Ser Lys Thr Ile Glu Ser Leu Lys Thr Ala Leu Arg Thr Lys  
130 135 140

Pro Met Arg Phe Val Thr Arg Phe Ile Asp Leu Asp Gly Leu Ser Cys  
145 150 155 160

Ile Leu Asn Phe Leu Lys Thr Met Asp Tyr Glu Thr Ser Glu Ser Arg  
165 170 175

Ile His Thr Ser Leu Ile Gly Cys Ile Lys Ala Leu Met Asn Asn Ser  
180 185 190

Gln Gly Arg Ala His Val Leu Ala His Ser Glu Ser Ile Asn Val Ile  
195 200 205

Ala Gln Ser Leu Ser Thr Glu Asn Ile Lys Thr Lys Val Ala Val Leu  
210 215 220

Glu Ile Leu Gly Ala Val Cys Leu Val Pro Gly Gly His Lys Lys Val  
225 230 235 240

Leu Gln Ala Met Leu His Tyr Gln Lys Tyr Ala Ser Glu Arg Thr Arg  
245 250 255

Phe Gln Thr Leu Ile Asn Asp Leu Asp Lys Ser Thr Gly Arg Tyr Arg  
260 265 270

Asp Glu Val Ser Leu Lys Thr Ala Ile Met Ser Phe Ile Asn Ala Val  
275 280 285

Leu Ser Gln Gly Ala Gly Val Glu Ser Leu Asp Phe Arg Leu His Leu  
290 295 300

Arg Tyr Glu Phe Leu Met Leu Gly Ile Gln Pro Val Ile Asp Lys Leu  
305 310 315 320

Arg Glu His Glu Asn Ser Thr Leu Asp Arg His Leu Asp Phe Phe Glu  
325 330 335

Met Leu Arg Asn Glu Asp Glu Leu Glu Phe Ala Lys Arg Phe Glu Leu  
340 345 350

Val His Ile Asp Thr Lys Ser Ala Thr Gln Met Phe Glu Leu Thr Arg  
355 360 365

Lys Arg Leu Thr His Ser Glu Ala Tyr Pro His Phe Met Ser Ile Leu  
370 375 380

His His Cys Leu Gln Met Pro Tyr Lys Arg Ser Gly Asn Thr Val Gln  
385 390 395 400

Tyr Trp Leu Leu Leu Asp Arg Ile Ile Gln Gln Ile Val Ile Gln Asn  
405 410 415

Asp Lys Gly Gln Asp Pro Asp Ser Thr Pro Leu Glu Asn Phe Asn Ile  
420 425 430

Lys Asn Val Val Arg Met Leu Val Asn Glu Asn Glu Val Lys Gln Trp  
435 440 445

Lys Glu Gln Ala Glu Lys Met Arg Lys Glu His Asn Glu Leu Gln Gln  
450 455 460

Lys Leu Glu Lys Lys Glu Arg Glu Cys Asp Ala Lys Thr Gln Glu Lys  
465 470 475 480

Glu Glu Met Met Gln Thr Leu Asn Lys Met Lys Glu Lys Leu Glu Lys  
485 490 495

Glu Thr Thr Glu His Lys Gln Val Lys Gln Gln Val Ala Asp Leu Thr  
 500 505 510

Ala Gln Leu His Glu Leu Ser Arg Arg Ala Val Cys Ala Ser Ile Pro  
 515 520 525

Gly Gly Pro Ser Pro Gly Ala Pro Gly Gly Pro Phe Pro Ser Ser Val  
 530 535 540

Pro Gly Ser Leu Leu Pro Pro Pro Pro Pro Pro Leu Pro Gly Gly  
 545 550 555 560

Met Leu Pro Pro Pro Pro Pro Pro Leu Pro Pro Gly Gly Pro Pro Pro  
 565 570 575

Pro Pro Gly Pro Pro Pro Leu Gly Ala Ile Met Pro Pro Pro Gly Ala  
 580 585 590

Pro Met Gly Leu Ala Leu Lys Lys Lys Ser Ile Pro Gln Pro Thr Asn  
 595 600 605

Ala Leu Lys Ser Phe Asn Trp Ser Lys Leu Pro Glu Asn Lys Leu Glu  
 610 615 620

Gly Thr Val Trp Thr Glu Ile Asp Asp Thr Lys Val Phe Lys Ile Leu  
 625 630 635 640

Asp Leu Glu Asp Leu Glu Arg Thr Phe Ser Ala Tyr Gln Arg Gln Gln  
 645 650 655

Asp Phe Phe Val Asn Ser Asn Ser Lys Gln Lys Glu Ala Asp Ala Ile  
 660 665 670

Asp Asp Thr Leu Ser Ser Lys Leu Lys Val Lys Glu Leu Ser Val Ile  
 675 680 685

Asp Gly Arg Arg Ala Gln Asn Cys Asn Ile Leu Leu Ser Arg Leu Lys  
 690 695 700

Leu Ser Asn Asp Glu Ile Lys Arg Ala Ile Leu Thr Met Asp Glu Gln  
 705 710 715 720

Glu Asp Leu Pro Lys Asp Met Leu Glu Gln Leu Leu Lys Phe Val Pro  
 725 730 735

Glu Lys Ser Asp Ile Asp Leu Leu Glu Gly His Lys His Glu Leu Asp  
 740 745 750

Arg Met Ala Lys Ala Asp Arg Phe Leu Phe Glu Met Ser Arg Ile Asn

755

760

765

His Tyr Gln Gln Arg Leu Gln Ser Leu Tyr Phe Lys Lys Lys Phe Ala  
 770 775 780

Glu Arg Val Ala Glu Val Lys Pro Lys Val Glu Ala Ile Arg Ser Gly  
 785 790 795 800

Ser Glu Glu Val Phe Arg Ser Gly Ala Leu Lys Gln Leu Leu Glu Val  
 805 810 815

Val Leu Ala Phe Gly Asn Tyr Met Asn Lys Gly Gln Arg Gly Asn Ala  
 820 825 830

Tyr Gly Phe Lys Ile Ser Ser Leu Asn Lys Ile Ala Asp Thr Lys Ser  
 835 840 845

Ser Ile Asp Lys Asn Ile Thr Leu Leu His Tyr Leu Ile Thr Ile Val  
 850 855 860

Glu Asn Lys Tyr Pro Ser Val Leu Asn Leu Asn Glu Glu Leu Arg Asp  
 865 870 875 880

Ile Pro Gln Ala Ala Lys Val Asn Met Thr Glu Leu Asp Lys Glu Ile  
 885 890 895

Ser Thr Leu Arg Ser Gly Leu Lys Ala Val Glu Thr Glu Leu Glu Tyr  
 900 905 910

Gln Lys Ser Gln Pro Pro Gln Pro Gly Asp Lys Phe Val Ser Val Val  
 915 920 925

Ser Gln Phe Ile Thr Val Ala Ser Phe Ser Phe Ser Asp Val Glu Asp  
 930 935 940

Leu Leu Ala Glu Ala Lys Asp Leu Phe Thr Lys Ala Val Lys His Phe  
 945 950 955 960

Gly Glu Glu Ala Gly Lys Ile Gln Pro Asp Glu Phe Phe Gly Ile Phe  
 965 970 975

Asp Gln Phe Leu Gln Ala Val Ser Glu Ala Lys Gln Glu Asn Glu Asn  
 980 985 990

Met Arg Lys Lys Lys Glu Glu Glu Glu Arg Arg Ala Arg Met Glu Ala  
 995 1000 1005

Gln Leu Lys Glu Gln Arg Glu Arg Glu Arg Lys Met Arg Lys Ala  
 1010 1015 1020

Lys Glu Asn Ser Glu Glu Ser Gly Glu Phe Asp Asp Leu Val Ser

1025                      1030                      1035

Ala Leu Arg Ser Gly Glu Val Phe Asp Lys Asp Leu Ser Lys Leu  
1040                      1045                      1050

Lys Arg Asn Arg Lys Arg Ile Thr Asn Gln Met Thr Asp Ser Ser  
1055                      1060                      1065

Arg Glu Arg Pro Ile Thr Lys Leu Asn Phe  
1070                      1075

&lt;210&gt; 67

&lt;211&gt; 1096

&lt;212&gt; DNA

&lt;213&gt; NM\_002122 MHC class II DQ alpha 1, HLA-DQA1

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(765)

&lt;223&gt;

<400> 67  
atg atc cta aac aaa gct ctg ctg ctg ggg gcc ctc gct ctg acc acc 48  
Met Ile Leu Asn Lys Ala Leu Leu Leu Gly Ala Leu Ala Leu Thr Thr  
1 5 10 15

gtg atg agc ccc tgt gga ggt gaa gac att gtg gct gac cac gtt gcc 96  
Val Met Ser Pro Cys Gly Gly Glu Asp Ile Val Ala Asp His Val Ala  
20 25 30

tct tgt ggt gta aac ttg tac cag ttt tac ggt ccc tct ggc cag tac 144  
Ser Cys Gly Val Asn Leu Tyr Gln Phe Tyr Gly Pro Ser Gly Gln Tyr  
35 40 45

acc cat gaa ttt gat gga gat gag cag ttc tac gtg gac ctg gag agg 192  
Thr His Glu Phe Asp Gly Asp Glu Gln Phe Tyr Val Asp Leu Glu Arg  
50 55 60

aag gag act gcc tgg cgg tgg cct gag ttc agc aaa ttt gga ggt ttt 240  
Lys Glu Thr Ala Trp Arg Trp Pro Glu Phe Ser Lys Phe Gly Gly Phe  
65 70 75 80

gac ccg cag ggt gca ctg aga aac atg gct gtg gca aaa cac aac ttg 288  
Asp Pro Gln Gly Ala Leu Arg Asn Met Ala Val Ala Lys His Asn Leu  
85 90 95

aac atc atg att aaa cgc tac aac tct acc gct gct acc aat gag gtt 336  
Asn Ile Met Ile Lys Arg Tyr Asn Ser Thr Ala Ala Thr Asn Glu Val  
100 105 110

cct gag gtc aca gtg ttt tcc aag tct ccc gtg aca ctg ggt cag ccc 384  
Pro Glu Val Thr Val Phe Ser Lys Ser Pro Val Thr Leu Gly Gln Pro  
115 120 125

aac acc ctc att tgt ctt gtg gac aac atc ttt cct cct gtg gtc aac 432  
 Asn Thr Leu Ile Cys Leu Val Asp Asn Ile Phe Pro Pro Val Val Asn  
 130 135 140  
 atc aca tgg ctg agc aat ggg cag tca gtc aca gaa ggt gtt tct gag 480  
 Ile Thr Trp Leu Ser Asn Gly Gln Ser Val Thr Glu Gly Val Ser Glu  
 145 150 155 160  
 acc agc ttc ctc tcc aag agt gat cat tcc ttc ttc aag atc agt tac 528  
 Thr Ser Phe Leu Ser Lys Ser Asp His Ser Phe Phe Lys Ile Ser Tyr  
 165 170 175  
 ctc acc ttc ctc cct tct gct gat gag att tat gac tgc aag gtg gag 576  
 Leu Thr Phe Leu Pro Ser Ala Asp Glu Ile Tyr Asp Cys Lys Val Glu  
 180 185 190  
 cac tgg ggc ctg gac cag cct ctt ctg aaa cac tgg gag cct gag att 624  
 His Trp Gly Leu Asp Gln Pro Leu Leu Lys His Trp Glu Pro Glu Ile  
 195 200 205  
 cca gcc cct atg tca gag ctc aca gag act gtg gtc tgt gcc ctg ggg 672  
 Pro Ala Pro Met Ser Glu Leu Thr Glu Thr Val Val Cys Ala Leu Gly  
 210 215 220  
 ttg tct gtg ggc ctc atg ggc att gtg gtg ggc act gtc ttc atc atc 720  
 Leu Ser Val Gly Leu Met Gly Ile Val Val Gly Thr Val Phe Ile Ile  
 225 230 235 240  
 caa ggc ctg cgt tca gtt ggt gct tcc aga cac caa ggg cca ttg 765  
 Gln Gly Leu Arg Ser Val Gly Ala Ser Arg His Gln Gly Pro Leu  
 245 250 255  
 tgaatcccat cctggaagg aagggtgcatc gccatctaca ggagcagaag aatggacttg 825  
 ctaaatgacc tagcactatt ctctggcccg atttatcata tcccttttct cctccaaata 885  
 tttctcctct caccttttct ctgggactta agctgctata tcccctcaga gctcacaaat 945  
 gcctttacat tctttccctg acctcctgat ttttttttct ttttttcaaa tgttacctac 1005  
 aatacatgcc tggggtaagc caccgggcta cctaattcct cagtaacctc catctaaaat 1065  
 ctccaaggaa gcaataaatt ccttttatga g 1096

&lt;210&gt; 68

&lt;211&gt; 255

&lt;212&gt; PRT

&lt;213&gt; NM\_002122 MHC class II DQ alpha 1, HLA-DQA1

&lt;400&gt; 68

Met Ile Leu Asn Lys Ala Leu Leu Leu Gly Ala Leu Ala Leu Thr Thr  
 1 5 10 15

Val Met Ser Pro Cys Gly Gly Glu Asp Ile Val Ala Asp His Val Ala  
 20 25 30

Ser Cys Gly Val Asn Leu Tyr Gln Phe Tyr Gly Pro Ser Gly Gln Tyr  
 35 40 45

Thr His Glu Phe Asp Gly Asp Glu Gln Phe Tyr Val Asp Leu Glu Arg  
 50 55 60  
 Lys Glu Thr Ala Trp Arg Trp Pro Glu Phe Ser Lys Phe Gly Gly Phe  
 65 70 75 80  
 Asp Pro Gln Gly Ala Leu Arg Asn Met Ala Val Ala Lys His Asn Leu  
 85 90 95  
 Asn Ile Met Ile Lys Arg Tyr Asn Ser Thr Ala Ala Thr Asn Glu Val  
 100 105 110  
 Pro Glu Val Thr Val Phe Ser Lys Ser Pro Val Thr Leu Gly Gln Pro  
 115 120 125  
 Asn Thr Leu Ile Cys Leu Val Asp Asn Ile Phe Pro Pro Val Val Asn  
 130 135 140  
 Ile Thr Trp Leu Ser Asn Gly Gln Ser Val Thr Glu Gly Val Ser Glu  
 145 150 155 160  
 Thr Ser Phe Leu Ser Lys Ser Asp His Ser Phe Phe Lys Ile Ser Tyr  
 165 170 175  
 Leu Thr Phe Leu Pro Ser Ala Asp Glu Ile Tyr Asp Cys Lys Val Glu  
 180 185 190  
 His Trp Gly Leu Asp Gln Pro Leu Leu Lys His Trp Glu Pro Glu Ile  
 195 200 205  
 Pro Ala Pro Met Ser Glu Leu Thr Glu Thr Val Val Cys Ala Leu Gly  
 210 215 220  
 Leu Ser Val Gly Leu Met Gly Ile Val Val Gly Thr Val Phe Ile Ile  
 225 230 235 240  
 Gln Gly Leu Arg Ser Val Gly Ala Ser Arg His Gln Gly Pro Leu  
 245 250 255

&lt;210&gt; 69

&lt;211&gt; 2820

&lt;212&gt; DNA

&lt;213&gt; NM\_003014 SFRP4, secreted frizzled-related protein 4

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (238) .. (1275)

&lt;223&gt;

&lt;400&gt; 69

```

ggcgggttcg cgccccgaag gctgagagct ggcgctgctc gtgccctgtg tgccagacgg      60
cggagctccg cggccggacc ccgcggcccc gctttgctgc cgactggagt ttgggggaag      120
aaactctcct gcgccccaga agatttcttc ctccggcgaag ggacagcgaa agatgagggt      180
ggcaggaaga gaaggcgctt tctgtctgcc ggggtcgag cgcgagaggg cagtgcc          237
atg ttc ctc tcc atc cta gtg gcg ctg tgc ctg tgg ctg cac ctg gcg          285
Met Phe Leu Ser Ile Leu Val Ala Leu Cys Leu Trp Leu His Leu Ala
1           5              10              15

ctg ggc gtg cgc ggc gcg ccc tgc gag gcg gtg cgc atc cct atg tgc          333
Leu Gly Val Arg Gly Ala Pro Cys Glu Ala Val Arg Ile Pro Met Cys
                20              25              30

cgg cac atg ccc tgg aac atc acg cgg atg ccc aac cac ctg cac cac          381
Arg His Met Pro Trp Asn Ile Thr Arg Met Pro Asn His Leu His His
                35              40              45

agc acg cag gag aac gcc atc ctg gcc atc gag cag tac gag gag ctg          429
Ser Thr Gln Glu Asn Ala Ile Leu Ala Ile Glu Gln Tyr Glu Glu Leu
                50              55              60

gtg gac gtg aac tgc agc gcc gtg ctg cgc ttc ttc ttc tgt gcc atg          477
Val Asp Val Asn Cys Ser Ala Val Leu Arg Phe Phe Phe Cys Ala Met
65              70              75              80

tac gcg ccc att tgc acc ctg gag ttc ctg cac gac cct atc aag ccg          525
Tyr Ala Pro Ile Cys Thr Leu Glu Phe Leu His Asp Pro Ile Lys Pro
                85              90              95

tgc aag tcg gtg tgc caa cgc gcg cgc gac gac tgc gag ccc ctc atg          573
Cys Lys Ser Val Cys Gln Arg Ala Arg Asp Asp Cys Glu Pro Leu Met
                100             105             110

aag atg tac aac cac agc tgg ccc gaa agc ctg gcc tgc gac gag ctg          621
Lys Met Tyr Asn His Ser Trp Pro Glu Ser Leu Ala Cys Asp Glu Leu
                115             120             125

cct gtc tat gac cgt ggc gtg tgc att tcg cct gaa gcc atc gtc acg          669
Pro Val Tyr Asp Arg Gly Val Cys Ile Ser Pro Glu Ala Ile Val Thr
                130             135             140

gac ctc ccg gag gat gtt aag tgg ata gac atc aca cca gac atg atg          717
Asp Leu Pro Glu Asp Val Lys Trp Ile Asp Ile Thr Pro Asp Met Met
145             150             155             160

gta cag gaa agg cct ctt gat gtt gac tgt aaa cgc cta agc ccc gat          765
Val Gln Glu Arg Pro Leu Asp Val Asp Cys Lys Arg Leu Ser Pro Asp
                165             170             175

cgg tgc aag tgt aaa aag gtg aag cca act ttg gca acg tat ctc agc          813
Arg Cys Lys Cys Lys Lys Val Lys Pro Thr Leu Ala Thr Tyr Leu Ser
                180             185             190

aaa aac tac agc tat gtt att cat gcc aaa ata aaa gct gtg cag agg          861
Lys Asn Tyr Ser Tyr Val Ile His Ala Lys Ile Lys Ala Val Gln Arg
                195             200             205

agt ggc tgc aat gag gtc aca acg gtg gtg gat gta aaa gag atc ttc          909

```





gacagttggg atactttaat cagaaaaaaa gaacttattt gcagcatttt atcaacaaat 2375  
 ttcataattg tggacaattg gaggcattta ttttaaaaaa caattttatt ggccttttgc 2435  
 taacacagta agcatgtatt ttataaggca ttcaataaat gcacaacgcc caaaggaaat 2495  
 aaaatcctat ctaatcctac tctccactac acagaggtaa tcactattag tattttggca 2555  
 tattattctc caggtgtttg cttatgcact tataaaatga ttgaacaaa taaaactagg 2615  
 aacctgtata catgtgtttc ataacctgcc tcctttgctt ggcctttat tgagataagt 2675  
 tttcctgtca agaaagcaga aaccatctca tttctaacag ctgtgttata ttccatagta 2735  
 tgcattactc aacaaactgt tgtgctattg gatacttagg tggtttcttc actgacaata 2795  
 ctgaataaac atctcaccgg aattc 2820

<210> 70

<211> 346

<212> PRT

<213> NM\_003014 SFRP4, secreted frizzled-related protein 4

<400> 70

Met Phe Leu Ser Ile Leu Val Ala Leu Cys Leu Trp Leu His Leu Ala  
 1 5 10 15

Leu Gly Val Arg Gly Ala Pro Cys Glu Ala Val Arg Ile Pro Met Cys  
 20 25 30

Arg His Met Pro Trp Asn Ile Thr Arg Met Pro Asn His Leu His His  
 35 40 45

Ser Thr Gln Glu Asn Ala Ile Leu Ala Ile Glu Gln Tyr Glu Glu Leu  
 50 55 60

Val Asp Val Asn Cys Ser Ala Val Leu Arg Phe Phe Phe Cys Ala Met  
 65 70 75 80

Tyr Ala Pro Ile Cys Thr Leu Glu Phe Leu His Asp Pro Ile Lys Pro  
 85 90 95

Cys Lys Ser Val Cys Gln Arg Ala Arg Asp Asp Cys Glu Pro Leu Met  
 100 105 110

Lys Met Tyr Asn His Ser Trp Pro Glu Ser Leu Ala Cys Asp Glu Leu  
 115 120 125

Pro Val Tyr Asp Arg Gly Val Cys Ile Ser Pro Glu Ala Ile Val Thr  
 130 135 140

Asp Leu Pro Glu Asp Val Lys Trp Ile Asp Ile Thr Pro Asp Met Met  
145 150 155 160

Val Gln Glu Arg Pro Leu Asp Val Asp Cys Lys Arg Leu Ser Pro Asp  
165 170 175

Arg Cys Lys Cys Lys Lys Val Lys Pro Thr Leu Ala Thr Tyr Leu Ser  
180 185 190

Lys Asn Tyr Ser Tyr Val Ile His Ala Lys Ile Lys Ala Val Gln Arg  
195 200 205

Ser Gly Cys Asn Glu Val Thr Thr Val Val Asp Val Lys Glu Ile Phe  
210 215 220

Lys Ser Ser Ser Pro Ile Pro Arg Thr Gln Val Pro Leu Ile Thr Asn  
225 230 235 240

Ser Ser Cys Gln Cys Pro His Ile Leu Pro His Gln Asp Val Leu Ile  
245 250 255

Met Cys Tyr Glu Trp Arg Ser Arg Met Met Leu Leu Glu Asn Cys Leu  
260 265 270

Val Glu Lys Trp Arg Asp Gln Leu Ser Lys Arg Ser Ile Gln Trp Glu  
275 280 285

Glu Arg Leu Gln Glu Gln Arg Arg Thr Val Gln Asp Lys Lys Lys Thr  
290 295 300

Ala Gly Arg Thr Ser Arg Ser Asn Pro Pro Lys Pro Lys Gly Lys Pro  
305 310 315 320

Pro Ala Pro Lys Pro Ala Ser Pro Lys Lys Asn Ile Lys Thr Arg Ser  
325 330 335

Ala Gln Lys Arg Thr Asn Pro Lys Arg Val  
340 345

<210> 71

<211> 1362

<212> DNA

<213> NM\_004039 annexin A2

<220>

<221> CDS

<222> (50)..(1066)

<400> 71

catttgggga cgctctcagc tctcggcgca cggccagct tccttcaa atg tct act 58  
 Met Ser Thr  
 1

gtt cac gaa atc ctg tgc aag ctc agc ttg gag ggt gat cac tct aca 106  
Val His Glu Ile Leu Cys Lys Leu Ser Leu Glu Gly Asp His Ser Thr  
5 10 15

ccc cca agt gca tat ggg tct gtc aaa gcc tat act aac ttt gat gct 154  
Pro Pro Ser Ala Tyr Gly Ser Val Lys Ala Tyr Thr Asn Phe Asp Ala  
20 25 30 35

gag cgg gat gct ttg aac att gaa aca gcc atc aag acc aaa ggt gtg . 202  
Glu Arg Asp Ala Leu Asn Ile Glu Thr Ala Ile Lys Thr Lys Gly Val  
40 45 50

gat gag gtc acc att gtc aac att ttg acc aac cgc agc aat gca cag 250  
Asp Glu Val Thr Ile Val Asn Ile Leu Thr Asn Arg Ser Asn Ala Gln  
55 60 65

aga cag gat att gcc ttc gcc tac cag aga agg acc aaa aag gaa ctt 298  
Arg Gln Asp Ile Ala Phe Ala Tyr Gln Arg Arg Thr Lys Lys Glu Leu  
70 75 80

gca tca gca ctg aag tca gcc tta tct ggc cac ctg gag acg gtg att 346  
Ala Ser Ala Leu Lys Ser Ala Leu Ser Gly His Leu Glu Thr Val Ile  
85 90 95

ttg ggc cta ttg aag aca cct gct cag tat gac gct tct gag cta aaa 394  
Leu Gly Leu Leu Lys Thr Pro Ala Gln Tyr Asp Ala Ser Glu Leu Lys  
100 105 110 115

gct tcc atg aag ggg ctg gga acc gac gag gac tct ctc att gag atc 442  
Ala Ser Met Lys Gly Leu Gly Thr Asp Glu Asp Ser Leu Ile Glu Ile  
120 125 130

atc tgc tcc aga acc aac cag gag ctg cag gaa att aac aga gtc tac 490  
Ile Cys Ser Arg Thr Asn Gln Glu Leu Gln Glu Ile Asn Arg Val Tyr  
135 140 145

aag gaa atg tac aag act gat ctg gag aag gac att att tcg gac aca 538  
Lys Glu Met Tyr Lys Thr Asp Leu Glu Lys Asp Ile Ile Ser Asp Thr  
150 155 160

tct ggt gac ttc cgc aag ctg atg gtt gcc ctg gca aag ggt aga aga 586  
Ser Gly Asp Phe Arg Lys Leu Met Val Ala Leu Ala Lys Gly Arg Arg  
165 170 175

gca gag gat ggc tct gtc att gat tat gaa ctg att gac caa gat gct 634  
Ala Glu Asp Gly Ser Val Ile Asp Tyr Glu Leu Ile Asp Gln Asp Ala  
180 185 190 195

cgg gat ctc tat gac gct gga gtg aag agg aaa gga act gat gtt ccc 682  
Arg Asp Leu Tyr Asp Ala Gly Val Lys Arg Lys Gly Thr Asp Val Pro  
200 205 210

aag tgg atc agc atc atg acc gag cgg agc gtg ccc cac ctc cag aaa 730  
Lys Trp Ile Ser Ile Met Thr Glu Arg Ser Val Pro His Leu Gln Lys  
215 220 225

gta ttt gat agg tac aag agt tac agc cct tat gac atg ttg gaa agc 778  
Val Phe Asp Arg Tyr Lys Ser Tyr Ser Pro Tyr Asp Met Leu Glu Ser  
230 235 240

atc agg aaa gag gtt aaa gga gac ctg gaa aat gct ttc ctg aac ctg 826  
 Ile Arg Lys Glu Val Lys Gly Asp Leu Glu Asn Ala Phe Leu Asn Leu  
 245 250 255

gtt cag tgc att cag aac aag ccc ctg tat ttt gct gat cgg ctg tat 874  
 Val Gln Cys Ile Gln Asn Lys Pro Leu Tyr Phe Ala Asp Arg Leu Tyr  
 260 265 270 275

gac tcc atg aag ggc aag ggg acg cga gat aag gtc ctg atc aga atc 922  
 Asp Ser Met Lys Gly Lys Gly Thr Arg Asp Lys Val Leu Ile Arg Ile  
 280 285 290

atg gtc tcc cgc agt gaa gtg gac atg ttg aaa att agg tct gaa ttc 970  
 Met Val Ser Arg Ser Glu Val Asp Met Leu Lys Ile Arg Ser Glu Phe  
 295 300 305

aag aga aag tac ggc aag tcc ctg tac tat tat atc cag caa gac act 1018  
 Lys Arg Lys Tyr Gly Lys Ser Leu Tyr Tyr Tyr Ile Gln Gln Asp Thr  
 310 315 320

aag ggc gac tac cag aaa gcg ctg ctg tac ctg tgt ggt gga gat gac 1066  
 Lys Gly Asp Tyr Gln Lys Ala Leu Leu Tyr Leu Cys Gly Gly Asp Asp  
 325 330 335

tgaagcccga cacggcctga gcgtccagaa atggtgctca ccatgcttcc agctaacagg 1126

tctagaaaac cagcttgcca ataacagtcc ccgtggccat ccctgtgagg gtgacgttag 1186

cattaccccc aacctcattt tagttgcta agcattgcct ggccttctct tctagtctct 1246

cctgtaagcc aaagaaatga acattccaag gagttggaag tgaagtctat gatgtgaaac 1306

actttgcctc ctgtgtactg tgtcataaac agatgaataa actgaatttg tactttt 1362

<210> 72

<211> 339

<212> PRT

<213> NM\_004039 annexin A2

<400> 72

Met Ser Thr Val His Glu Ile Leu Cys Lys Leu Ser Leu Glu Gly Asp  
 1 5 10 15

His Ser Thr Pro Pro Ser Ala Tyr Gly Ser Val Lys Ala Tyr Thr Asn  
 20 25 30

Phe Asp Ala Glu Arg Asp Ala Leu Asn Ile Glu Thr Ala Ile Lys Thr  
 35 40 45

Lys Gly Val Asp Glu Val Thr Ile Val Asn Ile Leu Thr Asn Arg Ser  
 50 55 60

Asn Ala Gln Arg Gln Asp Ile Ala Phe Ala Tyr Gln Arg Arg Thr Lys  
 65 70 75 80

Lys Glu Leu Ala Ser Ala Leu Lys Ser Ala Leu Ser Gly His Leu Glu  
85 90 95

Thr Val Ile Leu Gly Leu Leu Lys Thr Pro Ala Gln Tyr Asp Ala Ser  
100 105 110

Glu Leu Lys Ala Ser Met Lys Gly Leu Gly Thr Asp Glu Asp Ser Leu  
115 120 125

Ile Glu Ile Ile Cys Ser Arg Thr Asn Gln Glu Leu Gln Glu Ile Asn  
130 135 140

Arg Val Tyr Lys Glu Met Tyr Lys Thr Asp Leu Glu Lys Asp Ile Ile  
145 150 155 160

Ser Asp Thr Ser Gly Asp Phe Arg Lys Leu Met Val Ala Leu Ala Lys  
165 170 175

Gly Arg Arg Ala Glu Asp Gly Ser Val Ile Asp Tyr Glu Leu Ile Asp  
180 185 190

Gln Asp Ala Arg Asp Leu Tyr Asp Ala Gly Val Lys Arg Lys Gly Thr  
195 200 205

Asp Val Pro Lys Trp Ile Ser Ile Met Thr Glu Arg Ser Val Pro His  
210 215 220

Leu Gln Lys Val Phe Asp Arg Tyr Lys Ser Tyr Ser Pro Tyr Asp Met  
225 230 235 240

Leu Glu Ser Ile Arg Lys Glu Val Lys Gly Asp Leu Glu Asn Ala Phe  
245 250 255

Leu Asn Leu Val Gln Cys Ile Gln Asn Lys Pro Leu Tyr Phe Ala Asp  
260 265 270

Arg Leu Tyr Asp Ser Met Lys Gly Lys Gly Thr Arg Asp Lys Val Leu  
275 280 285

Ile Arg Ile Met Val Ser Arg Ser Glu Val Asp Met Leu Lys Ile Arg  
290 295 300

Ser Glu Phe Lys Arg Lys Tyr Gly Lys Ser Leu Tyr Tyr Tyr Ile Gln  
305 310 315 320

Gln Asp Thr Lys Gly Asp Tyr Gln Lys Ala Leu Leu Tyr Leu Cys Gly  
325 330 335

Gly Asp Asp

&lt;210&gt; 73

&lt;211&gt; 850

&lt;212&gt; DNA

&lt;213&gt; NM\_003955 SOCS3

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (107) .. (781)

&lt;223&gt;

&lt;400&gt; 73

```

gcgccttcct ctccgcagcc ccccgggatg cggtagcggc cgctgtgcgg aggccgcgaa      60
gcagctgcag ccgcgcgcgc gcagatccac gctggctccg tgcgcc atg gtc acc      115
                                   Met Val Thr
                                   1

cac agc aag ttt ccc gcc gcc ggg atg agc cgc ccc ctg gac acc agc      163
His Ser Lys Phe Pro Ala Ala Gly Met Ser Arg Pro Leu Asp Thr Ser
   5                                10                                15

ctg cgc ctc aag acc ttc agc tcc aag agc gag tac cag ctg gtg gtg      211
Leu Arg Leu Lys Thr Phe Ser Ser Lys Ser Glu Tyr Gln Leu Val Val
  20                                25                                30                                35

aac gca gtg cgc aag ctg cag gag agc ggc ttc tac tgg agc gca gtg      259
Asn Ala Val Arg Lys Leu Gln Glu Ser Gly Phe Tyr Trp Ser Ala Val
   40                                45                                50

acc ggc ggc gag gcg aac ctg ctg ctc agt gcc gag ccc gcc ggc acc      307
Thr Gly Gly Glu Ala Asn Leu Leu Leu Ser Ala Glu Pro Ala Gly Thr
   55                                60                                65

ttt ctg atc cgc gac agc tcg gac cag cgc cac ttc ttc acg ctc agc      355
Phe Leu Ile Arg Asp Ser Ser Asp Gln Arg His Phe Phe Thr Leu Ser
   70                                75                                80

gtc aag acc cag tct ggg acc aag aac ctg cgc atc cag tgt gag ggg      403
Val Lys Thr Gln Ser Gly Thr Lys Asn Leu Arg Ile Gln Cys Glu Gly
   85                                90                                95

ggc agc ttc tct ctg cag agc gat ccc cgg agc acg cag ccc gtg ccc      451
Gly Ser Phe Ser Leu Gln Ser Asp Pro Arg Ser Thr Gln Pro Val Pro
  100                                105                                110                                115

cgc ttc gac tgc gtg ctc aag ctg gtg tac cac tac atg ccg ccc cct      499
Arg Phe Asp Cys Val Leu Lys Leu Val Tyr His Tyr Met Pro Pro Pro
   120                                125                                130

gga gcc ccc tcc ttc ccc tcg cca cct act gaa ccc tcc tcc gag gtg      547
Gly Ala Pro Ser Phe Pro Ser Pro Thr Glu Pro Ser Ser Glu Val
   135                                140                                145

ccc gag cag ccg tct gcc cag cca ctc cct ggg agt ccc ccc aga aga      595
Pro Glu Gln Pro Ser Ala Gln Pro Leu Pro Gly Ser Pro Pro Arg Arg
   150                                155                                160

gcc tat tac atc tac tcc ggg ggc gag aag atc ccc ctg gtg ttg agc      643

```

Ala Tyr Tyr Ile Tyr Ser Gly Gly Glu Lys Ile Pro Leu Val Leu Ser  
 165 170 175

cgg ccc ctc tcc tcc aac gtg gcc act ctt cag cat ctc tgt cgg aag 691  
 Arg Pro Leu Ser Ser Asn Val Ala Thr Leu Gln His Leu Cys Arg Lys  
 180 185 190 195

acc gtc aac ggc cac ctg gac tcc tat gag aaa gtc acc cag ctg ccg 739  
 Thr Val Asn Gly His Leu Asp Ser Tyr Glu Lys Val Thr Gln Leu Pro  
 200 205 210

ggg ccc att cgg gag ttc ctg gac cag tac gat gcc ccg ctt 781  
 Gly Pro Ile Arg Glu Phe Leu Asp Gln Tyr Asp Ala Pro Leu  
 215 220 225

taaggggtaa agggcgcaaa gggcatgggt cgggagaggg gacgcaggcc cctctcctcc 841  
 gtggcacat 850

&lt;210&gt; 74

&lt;211&gt; 225

&lt;212&gt; PRT

&lt;213&gt; NM\_003955 SOCS3

&lt;400&gt; 74

Met Val Thr His Ser Lys Phe Pro Ala Ala Gly Met Ser Arg Pro Leu  
 1 5 10 15

Asp Thr Ser Leu Arg Leu Lys Thr Phe Ser Ser Lys Ser Glu Tyr Gln  
 20 25 30

Leu Val Val Asn Ala Val Arg Lys Leu Gln Glu Ser Gly Phe Tyr Trp  
 35 40 45

Ser Ala Val Thr Gly Gly Glu Ala Asn Leu Leu Leu Ser Ala Glu Pro  
 50 55 60

Ala Gly Thr Phe Leu Ile Arg Asp Ser Ser Asp Gln Arg His Phe Phe  
 65 70 75 80

Thr Leu Ser Val Lys Thr Gln Ser Gly Thr Lys Asn Leu Arg Ile Gln  
 85 90 95

Cys Glu Gly Gly Ser Phe Ser Leu Gln Ser Asp Pro Arg Ser Thr Gln  
 100 105 110

Pro Val Pro Arg Phe Asp Cys Val Leu Lys Leu Val Tyr His Tyr Met  
 115 120 125

Pro Pro Pro Gly Ala Pro Ser Phe Pro Ser Pro Pro Thr Glu Pro Ser  
 130 135 140



Ser Glu Val Pro Glu Gln Pro Ser Ala Gln Pro Leu Pro Gly Ser Pro  
145 150 155 160

Pro Arg Arg Ala Tyr Tyr Ile Tyr Ser Gly Gly Glu Lys Ile Pro Leu  
165 170 175

Val Leu Ser Arg Pro Leu Ser Ser Asn Val Ala Thr Leu Gln His Leu  
180 185 190

Cys Arg Lys Thr Val Asn Gly His Leu Asp Ser Tyr Glu Lys Val Thr  
195 200 205

Gln Leu Pro Gly Pro Ile Arg Glu Phe Leu Asp Gln Tyr Asp Ala Pro  
210 215 220

Leu  
225

<210> 75

<211> 369

<212> DNA

<213> NM\_000331 SAA1, serum amyloid A1

<220>

<221> CDS

<222> (1)..(366)

<223>

<400> 75

atg aag ctt ctc acg ggc ctg gtt ttc tgc tcc ttg gtc ctg ggt gtc 48  
Met Lys Leu Leu Thr Gly Leu Val Phe Cys Ser Leu Val Leu Gly Val  
1 5 10 15

agc agc cga agc ttc ttt tcg ttc ctt ggc gag gct ttt gat ggg gct 96  
Ser Ser Arg Ser Phe Phe Ser Phe Leu Gly Glu Ala Phe Asp Gly Ala  
20 25 30

cgg gac atg tgg aga gcc tac tct gac atg aga gaa gcc aat tac atc 144  
Arg Asp Met Trp Arg Ala Tyr Ser Asp Met Arg Glu Ala Asn Tyr Ile  
35 40 45

ggc tca gac aaa tac ttc cat gct cgg ggg aac tat gat gct gcc aaa 192  
Gly Ser Asp Lys Tyr Phe His Ala Arg Gly Asn Tyr Asp Ala Ala Lys  
50 55 60

agg gga cct ggg ggt gtc tgg gct gca gaa gcg atc agc gat gcc aga 240  
Arg Gly Pro Gly Gly Val Trp Ala Ala Glu Ala Ile Ser Asp Ala Arg  
65 70 75 80

gag aat atc cag aga ttc ttt ggc cat ggt gcg gag gac tcg ctg gct 288  
Glu Asn Ile Gln Arg Phe Phe Gly His Gly Ala Glu Asp Ser Leu Ala  
85 90 95

gat cag gct gcc aat gaa tgg ggc agg agt ggc aaa gac ccc aat cac 336  
Asp Gln Ala Ala Asn Glu Trp Gly Arg Ser Gly Lys Asp Pro Asn His  
100 105 110

ttc cga cct gct ggc ctg cct gag aaa tac tga 369  
Phe Arg Pro Ala Gly Leu Pro Glu Lys Tyr  
115 120

<210> 76

<211> 122

<212> PRT

<213> NM\_000331 SAA1, serum amyloid A1

<400> 76

Met Lys Leu Leu Thr Gly Leu Val Phe Cys Ser Leu Val Leu Gly Val  
1 5 10 15

Ser Ser Arg Ser Phe Phe Ser Phe Leu Gly Glu Ala Phe Asp Gly Ala  
20 25 30

Arg Asp Met Trp Arg Ala Tyr Ser Asp Met Arg Glu Ala Asn Tyr Ile  
35 40 45

Gly Ser Asp Lys Tyr Phe His Ala Arg Gly Asn Tyr Asp Ala Ala Lys  
50 55 60

Arg Gly Pro Gly Gly Val Trp Ala Ala Glu Ala Ile Ser Asp Ala Arg  
65 70 75 80

Glu Asn Ile Gln Arg Phe Phe Gly His Gly Ala Glu Asp Ser Leu Ala  
85 90 95

Asp Gln Ala Ala Asn Glu Trp Gly Arg Ser Gly Lys Asp Pro Asn His  
100 105 110

Phe Arg Pro Ala Gly Leu Pro Glu Lys Tyr  
115 120

<210> 77

<211> 895

<212> DNA

<213> NM\_014059 RGC32

<220>

<221> CDS

&lt;222&gt; (147) .. (497)

&lt;223&gt;

&lt;400&gt; 77

gcgggccgcgt cgaccggcgc ggctggagcg cagcgccgaa gggactggca gggctgaagt 60  
gtgcgggaca gcaagccccc gaatagcccc ggctgccacc tcgcaggacc caaggccacg 120  
cgcgccgggc ccagctgagc cgctc atg aag ccg ccc gcg gag gac ctg tcg 173  
Met Lys Pro Pro Ala Glu Asp Leu Ser  
1 5  
gac gcg ctg tgc gag ttt gac gcg gtg ctg gcc gac ttc gcg tcg ccc 221  
Asp Ala Leu Cys Glu Phe Asp Ala Val Leu Ala Asp Phe Ala Ser Pro  
10 15 20 25  
ttc cac gag cgc cac ttc cac tac gag gag cac ctg gag cgc atg aag 269  
Phe His Glu Arg His Phe His Tyr Glu Glu His Leu Glu Arg Met Lys  
30 35 40  
cgg cgc agc agc gcc agt gtc agc gac agc agc ggc ttc agc gac tcg 317  
Arg Arg Ser Ser Ala Ser Val Ser Asp Ser Ser Gly Phe Ser Asp Ser  
45 50 55  
gag agt gca gat tca ctt tat agg aac agc ttc agc ttc agt gat gaa 365  
Glu Ser Ala Asp Ser Leu Tyr Arg Asn Ser Phe Ser Phe Ser Asp Glu  
60 65 70  
aaa ctg aat tct cca aca gac tct acc cca gct ctt ctc tct gcc act 413  
Lys Leu Asn Ser Pro Thr Asp Ser Thr Pro Ala Leu Leu Ser Ala Thr  
75 80 85  
gtc act cct cag aaa gct aaa tta gga gac aca aaa gag cta gaa gcc 461  
Val Thr Pro Gln Lys Ala Lys Leu Gly Asp Thr Lys Glu Leu Glu Ala  
90 95 100 105  
ttc att gct gat ctt gac aaa act tta gca agt atg tgaacaaga 507  
Phe Ile Ala Asp Leu Asp Lys Thr Leu Ala Ser Met  
110 115  
agttctgggt cctttcatca taaggagaa gcttcagaaa gttccgagga cctgctaaaa 567  
tcagctacta gaatctgctg ccagagggga caaagacgtg cactcaacct tctaccaggc 627  
cactctcagg ctcaccttaa aatcagccct tgatcccat tctgggcaat ttagacagt 687  
aaactgactt tgtttacctg cttgcagcat attagaacag acgatccatg ctaatatgt 747  
attttctctt aaaacatagc tttcctgtaa tttaaagtgc ttttatgaaa atatttgtaa 807  
ttaattatat atagtggaa atagcagtaa gctttcccat tataatatat tttgtatac 867  
aaataaaatt tgaactgaac ctcgtgcc 895

&lt;210&gt; 78

&lt;211&gt; 117

&lt;212&gt; PRT

&lt;213&gt; NM\_014059 RGC32

&lt;400&gt; 78

Met Lys Pro Pro Ala Glu Asp Leu Ser Asp Ala Leu Cys Glu Phe Asp  
1 5 10 15

Ala Val Leu Ala Asp Phe Ala Ser Pro Phe His Glu Arg His Phe His  
20 25 30

Tyr Glu Glu His Leu Glu Arg Met Lys Arg Arg Ser Ser Ala Ser Val  
35 40 45

Ser Asp Ser Ser Gly Phe Ser Asp Ser Glu Ser Ala Asp Ser Leu Tyr  
50 55 60

Arg Asn Ser Phe Ser Phe Ser Asp Glu Lys Leu Asn Ser Pro Thr Asp  
65 70 75 80

Ser Thr Pro Ala Leu Leu Ser Ala Thr Val Thr Pro Gln Lys Ala Lys  
85 90 95

Leu Gly Asp Thr Lys Glu Leu Glu Ala Phe Ile Ala Asp Leu Asp Lys  
100 105 110

Thr Leu Ala Ser Met  
115

&lt;210&gt; 79

&lt;211&gt; 1564

&lt;212&gt; DNA

&lt;213&gt; NM\_018004 FLJ10134

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (314)..(1138)

&lt;223&gt;

&lt;400&gt; 79

gaagcacatc tggacagctg tgcggcctcc ttgcgggccg acgtcagccg agcacgtccc 60  
ccacgtcctc tccttctcgc cacttattat ttattcgttt tcccaaagaa gcgactaggg 120  
acccaagttt aaaaattcct ccccccactc aatgcgagac gtggccagat cccatccaac 180  
acacggttta attttcatgg ggctctggga tcaaaagaac agaaacagca acaacaaaag 240  
cccagccgct gtctgatttt aagctggcaa agtgggaaaa ataaagtgtt gagtaaacag 300  
accaagttgg atc atg ggg aat ttc aga ggt cat gcc ctc cct gga acc 349  
Met Gly Asn Phe Arg Gly His Ala Leu Pro Gly Thr  
1 5 10

|   |      |
|---|------|
| ttc ttt ttt att att ggt ctt tgg tgg tgt aca aag agt att ctg aag<br>Phe Phe Phe Ile Ile Gly Leu Trp Trp Cys Thr Lys Ser Ile Leu Lys<br>15 20 25        | 397  |
| tat atc tgc aaa aag caa aag cga acc tgc tat ctt ggt tcc aaa aca<br>Tyr Ile Cys Lys Lys Gln Lys Arg Thr Cys Tyr Leu Gly Ser Lys Thr<br>30 35 40        | 445  |
| tta ttc tat cga ttg gaa att ttg gag gga att aca ata gtt ggc atg<br>Leu Phe Tyr Arg Leu Glu Ile Leu Glu Gly Ile Thr Ile Val Gly Met<br>45 50 55 60     | 493  |
| gct tta act ggc atg gct ggg gag cag ttt att cct gga ggg ccc cat<br>Ala Leu Thr Gly Met Ala Gly Glu Gln Phe Ile Pro Gly Gly Pro His<br>65 70 75        | 541  |
| ctg atg tta tat gac tat aaa caa ggt cac tgg aat caa ctc ctg ggc<br>Leu Met Leu Tyr Asp Tyr Lys Gln Gly His Trp Asn Gln Leu Leu Gly<br>80 85 90        | 589  |
| tgg cat cat ttc acc atg tat ttc ttc ttt ggg ctg ttg ggt gtg gca<br>Trp His His Phe Thr Met Tyr Phe Phe Phe Gly Leu Leu Gly Val Ala<br>95 100 105      | 637  |
| gat atc tta tgt ttc acc atc agt tca ctt cct gtg tcc tta acc aag<br>Asp Ile Leu Cys Phe Thr Ile Ser Ser Leu Pro Val Ser Leu Thr Lys<br>110 115 120     | 685  |
| tta atg ttg tca aat gcc tta ttt gtg gag gcc ttt atc ttc tac aac<br>Leu Met Leu Ser Asn Ala Leu Phe Val Glu Ala Phe Ile Phe Tyr Asn<br>125 130 135 140 | 733  |
| cac act cat ggc cgg gaa atg ctg gac atc ttt gtg cac cag ctg ctg<br>His Thr His Gly Arg Glu Met Leu Asp Ile Phe Val His Gln Leu Leu<br>145 150 155     | 781  |
| gtt ttg gtc gtc ttt ctg aca ggc ctc gtt gcc ttc cta gag ttc ctt<br>Val Leu Val Val Phe Leu Thr Gly Leu Val Ala Phe Leu Glu Phe Leu<br>160 165 170     | 829  |
| gtt cgg aac aat gta ctt ctg gag cta ttg cgg tca agt ctc att ctg<br>Val Arg Asn Asn Val Leu Leu Glu Leu Leu Arg Ser Ser Leu Ile Leu<br>175 180 185     | 877  |
| ctt cag ggg agc tgg ttc ttt cag att gga ttt gtc ctg tat ccc ccc<br>Leu Gln Gly Ser Trp Phe Phe Gln Ile Gly Phe Val Leu Tyr Pro Pro<br>190 195 200     | 925  |
| agt gga ggt cct gca tgg gat ctg atg gat cat gaa aat att ttg ttt<br>Ser Gly Gly Pro Ala Trp Asp Leu Met Asp His Glu Asn Ile Leu Phe<br>205 210 215 220 | 973  |
| ctc acc ata tgc ttt tgt tgg cat tat gca gta acc att gtc atc gtt<br>Leu Thr Ile Cys Phe Cys Trp His Tyr Ala Val Thr Ile Val Ile Val<br>225 230 235     | 1021 |
| gga atg aat tat gct ttc att acc tgg ttg gtt aaa tct aga ctt aag<br>Gly Met Asn Tyr Ala Phe Ile Thr Trp Leu Val Lys Ser Arg Leu Lys<br>240 245 250     | 1069 |
| agg ctc tgc tcc tca gaa gtt gga ctt ctg aaa aat gct gaa cga gaa<br>Arg Leu Cys Ser Ser Glu Val Gly Leu Leu Lys Asn Ala Glu Arg Glu<br>255 260 265     | 1117 |
| caa gaa tca gaa gaa gaa atg tgactttgat gagcttcag tttttctaga<br>Gln Glu Ser Glu Glu Glu Met<br>270 275   | 1168 |

taaacctttt cttttttaca ttgttcttgg ttttgtttct cgatcttttg ttgggagaac 1228  
agctggctaa ggatgactct aagtgtactg ttgcatcttc caatttggtt aaagtatttg 1288  
aatttaaata ttttcttttt agctttgaaa atattttggg tgatactttc attttgcaca 1348  
tcatgcacat catggtattc aggggctaga gtgatttttt tccagattat ctaaagttag 1408  
atgcccacac tatgaaagaa atatttggtt tatttgcctt atagatatgc tcaaggttac 1468  
tgggcttgct actatttgta actccttgac catggaatta tacttgttta tcttgttgct 1528  
gcaatgagaa ataaatgaat gtatgtattt tgggtgc 1564

&lt;210&gt; 80

&lt;211&gt; 275

&lt;212&gt; PRT

&lt;213&gt; NM\_018004 FLJ10134

&lt;400&gt; 80

Met Gly Asn Phe Arg Gly His Ala Leu Pro Gly Thr Phe Phe Phe Ile  
1 5 10 15

Ile Gly Leu Trp Trp Cys Thr Lys Ser Ile Leu Lys Tyr Ile Cys Lys  
20 25 30

Lys Gln Lys Arg Thr Cys Tyr Leu Gly Ser Lys Thr Leu Phe Tyr Arg  
35 40 45

Leu Glu Ile Leu Glu Gly Ile Thr Ile Val Gly Met Ala Leu Thr Gly  
50 55 60

Met Ala Gly Glu Gln Phe Ile Pro Gly Gly Pro His Leu Met Leu Tyr  
65 70 75 80

Asp Tyr Lys Gln Gly His Trp Asn Gln Leu Leu Gly Trp His His Phe  
85 90 95

Thr Met Tyr Phe Phe Phe Gly Leu Leu Gly Val Ala Asp Ile Leu Cys  
100 105 110

Phe Thr Ile Ser Ser Leu Pro Val Ser Leu Thr Lys Leu Met Leu Ser  
115 120 125

Asn Ala Leu Phe Val Glu Ala Phe Ile Phe Tyr Asn His Thr His Gly  
130 135 140

Arg Glu Met Leu Asp Ile Phe Val His Gln Leu Leu Val Leu Val Val  
145 150 155 160

Phe Leu Thr Gly Leu Val Ala Phe Leu Glu Phe Leu Val Arg Asn Asn  
 165 170 175

Val Leu Leu Glu Leu Leu Arg Ser Ser Leu Ile Leu Leu Gln Gly Ser  
 180 185 190

Trp Phe Phe Gln Ile Gly Phe Val Leu Tyr Pro Pro Ser Gly Gly Pro  
 195 200 205

Ala Trp Asp Leu Met Asp His Glu Asn Ile Leu Phe Leu Thr Ile Cys  
 210 215 220

Phe Cys Trp His Tyr Ala Val Thr Ile Val Ile Val Gly Met Asn Tyr  
 225 230 235 240

Ala Phe Ile Thr Trp Leu Val Lys Ser Arg Leu Lys Arg Leu Cys Ser  
 245 250 255

Ser Glu Val Gly Leu Leu Lys Asn Ala Glu Arg Glu Gln Glu Ser Glu  
 260 265 270

Glu Glu Met  
 275

<210> 81

<211> 2311

<212> DNA

<213> NM\_004004 GJB2, connexin 26

<220>

<221> CDS

<222> (199)..(876)

<223>

<400> 81

gatttaatcc tatgacaaac taagttgggt ctgtcttcac ctgttttggt gaggttgtgt 60

aagagttggt gtttgctcag gaagagattt aagcatgctt gcttaccag actcagagaa 120

gtctccctgt tctgtcctag ctatgttctt gtgttggtg cattcgtctt ttccagagca 180

aaccgcccag agtagaag atg gat tgg ggc acg ctg cag acg atc ctg ggg 231  
 Met Asp Trp Gly Thr Leu Gln Thr Ile Leu Gly  
 1 5 10

ggt gtg aac aaa cac tcc acc agc att gga aag atc tgg ctc acc gtc 279  
 Gly Val Asn Lys His Ser Thr Ser Ile Gly Lys Ile Trp Leu Thr Val  
 15 20 25

ctc ttc att ttt cgc att atg atc ctc gtt gtg gct gca aag gag gtg 327

|  |      |
|--|------|
| Leu Phe Ile Phe Arg Ile Met Ile Leu Val Val Ala Ala Lys Glu Val    |      |
| 30 35 40   |      |
| tgg gga gat gag cag gcc gac ttt gtc tgc aac acc ctg cag cca ggc    | 375  |
| Trp Gly Asp Glu Gln Ala Asp Phe Val Cys Asn Thr Leu Gln Pro Gly    |      |
| 45 50 55   |      |
| tgc aag aac gtg tgc tac gat cac tac ttc ccc atc tcc cac atc cgg    | 423  |
| Cys Lys Asn Val Cys Tyr Asp His Tyr Phe Pro Ile Ser His Ile Arg    |      |
| 60 65 70 75  |      |
| cta tgg gcc ctg cag ctg atc ttc gtg tcc agc cca gcg ctc cta gtg    | 471  |
| Leu Trp Ala Leu Gln Leu Ile Phe Val Ser Ser Pro Ala Leu Leu Val    |      |
| 80 85 90   |      |
| gcc atg cac gtg gcc tac cgg aga cat gag aag aag agg aag ttc atc    | 519  |
| Ala Met His Val Ala Tyr Arg Arg His Glu Lys Lys Arg Lys Phe Ile    |      |
| 95 100 105   |      |
| aag ggg gag ata aag agt gaa ttt aag gac atc gag gag atc aaa acc    | 567  |
| Lys Gly Glu Ile Lys Ser Glu Phe Lys Asp Ile Glu Glu Ile Lys Thr    |      |
| 110 115 120  |      |
| cag aag gtc cgc atc gaa gcc tcc ctg tgg tgg acc tac aca agc agc    | 615  |
| Gln Lys Val Arg Ile Glu Gly Ser Leu Trp Trp Thr Tyr Thr Ser Ser    |      |
| 125 130 135  |      |
| atc ttc ttc cgg gtc atc ttc gaa gcc gcc ttc atg tac gtc ttc tat    | 663  |
| Ile Phe Phe Arg Val Ile Phe Glu Ala Ala Phe Met Tyr Val Phe Tyr    |      |
| 140 145 150 155  |      |
| gtc atg tac gac gcc ttc tcc atg cag cgg ctg gtg aag tgc aac gcc    | 711  |
| Val Met Tyr Asp Gly Phe Ser Met Gln Arg Leu Val Lys Cys Asn Ala    |      |
| 160 165 170  |      |
| tgg cct tgt ccc aac act gtg gac tgc ttt gtg tcc cgg ccc acg gag    | 759  |
| Trp Pro Cys Pro Asn Thr Val Asp Cys Phe Val Ser Arg Pro Thr Glu    |      |
| 175 180 185  |      |
| aag act gtc ttc aca gtg ttc atg att gca gtg tct gga att tgc atc    | 807  |
| Lys Thr Val Phe Thr Val Phe Met Ile Ala Val Ser Gly Ile Cys Ile    |      |
| 190 195 200  |      |
| ctg ctg aat gtc act gaa ttg tgt tat ttg cta att aga tat tgt tct    | 855  |
| Leu Leu Asn Val Thr Glu Leu Cys Tyr Leu Leu Ile Arg Tyr Cys Ser    |      |
| 205 210 215  |      |
| ggg aag tca aaa aag cca gtt taacgcattg cccagttggt agattaagaa       | 906  |
| Gly Lys Ser Lys Lys Pro Val  |      |
| 220 225  |      |
| atagacagca tgagagggat gaggcaaccc gtgctcagct gtcaaggctc agtcgccagc  | 966  |
| atttccaac acaaagattc tgaccttaaa tgcaaccatt tgaaacccct gtaggoccca   | 1026 |
| ggtgaaactc cagatgccac aatgagctct gctcccctaa agcctcaaaa caaaggccta  | 1086 |
| attctatgcc tgtcttaatt ttctttcact taagttagtt ccaactgagac cccaggctgt | 1146 |
| taggggttat tgggtgaagg tactttcata ttttaaacag aggatatcgg catttgtttc  | 1206 |
| tttctctgag gacaagagaa aaaagccagg ttccacagag gacacagaga aggtttgggt  | 1266 |
| gtcctcctgg ggttcttttt gccaaacttc cccacgttaa aggtgaacat tggttctttc  | 1326 |
| atttgctttg gaagttttta tctctaacag tggacaaagt taccagtgcc ttaaactctg  | 1386 |
| ttacactttt tggaagtga aactttgtag tatgataggt tattttgatg taaagatggt   | 1446 |



ctggatacca ttatatgttc cccctgttcc agaggctcag attgtaatat gtaaattgta 1506  
 tgtcattcgc tactatgatt taatttgaaa tatggctctt tgggtatgaa tactttgcag 1566  
 cacagctgag agaggctgtc tgttgtattc atttgtgtca tagcacctaa caacattgta 1626  
 gcctcaatcg agtgagacag actagaagtt cctagttggc ttatgatagc aaatggcctc 1686  
 atgtcaaata ttagatgtaa ttttgtgtaa gaaatacaga ctggatgtac caccaactac 1746  
 tacctgtaat gacaggcctg tccaacacat ctcccttttc catgctgtgg tagccagcat 1806  
 cggaaagaac gctgatttaa agaggtgagc ttgggaattt tattgacaca gtaccattta 1866  
 atggggagac aaaaatgggg gccaggggag ggagaagttt ctgtcgtaa aaacgagttt 1926  
 ggaaagactg gactctaaat tctgttgatt aaagatgagc tttgtctacc ttcaaaagtt 1986  
 tgtttggett acccccttca gcctccaatt ttttaagtga aaatataact aataacatgt 2046  
 gaaaagaata gaagctaagg tttagataaa tattgagcag atctatagga agattgaacc 2106  
 tgaatattgc cattatgctt gacatgggtt ccaaaaaatg gtactccaca tacttcagtg 2166  
 agggtaagta ttttctgtt gtcaagaata gcattgtaaa agcattttgt aataataaag 2226  
 aatagcttta atgatatgct tgtaactaaa ataattttgt aatgtatcaa atacatttaa 2286  
 aacattaaaa tataatctct ataat 2311

<210> 82

<211> 226

<212> PRT

<213> NM\_004004 GJB2, connexin 26

<400> 82

Met Asp Trp Gly Thr Leu Gln Thr Ile Leu Gly Gly Val Asn Lys His  
1 5 10 15

Ser Thr Ser Ile Gly Lys Ile Trp Leu Thr Val Leu Phe Ile Phe Arg  
20 25 30

Ile Met Ile Leu Val Val Ala Ala Lys Glu Val Trp Gly Asp Glu Gln  
35 40 45

Ala Asp Phe Val Cys Asn Thr Leu Gln Pro Gly Cys Lys Asn Val Cys  
50 55 60

Tyr Asp His Tyr Phe Pro Ile Ser His Ile Arg Leu Trp Ala Leu Gln  
65 70 75 80

Leu Ile Phe Val Ser Ser Pro Ala Leu Leu Val Ala Met His Val Ala  
85 90 95

Tyr Arg Arg His Glu Lys Lys Arg Lys Phe Ile Lys Gly Glu Ile Lys  
 100 105 110  
 Ser Glu Phe Lys Asp Ile Glu Glu Ile Lys Thr Gln Lys Val Arg Ile  
 115 120 125  
 Glu Gly Ser Leu Trp Trp Thr Tyr Thr Ser Ser Ile Phe Phe Arg Val  
 130 135 140  
 Ile Phe Glu Ala Ala Phe Met Tyr Val Phe Tyr Val Met Tyr Asp Gly  
 145 150 155 160  
 Phe Ser Met Gln Arg Leu Val Lys Cys Asn Ala Trp Pro Cys Pro Asn  
 165 170 175  
 Thr Val Asp Cys Phe Val Ser Arg Pro Thr Glu Lys Thr Val Phe Thr  
 180 185 190  
 Val Phe Met Ile Ala Val Ser Gly Ile Cys Ile Leu Leu Asn Val Thr  
 195 200 205  
 Glu Leu Cys Tyr Leu Leu Ile Arg Tyr Cys Ser Gly Lys Ser Lys Lys  
 210 215 220

Pro Val  
225

<210> 83

<211> 2389

<212> DNA

<213> NM\_002514 NOV1, nephroblastoma overexpressed gene

<220>

<221> CDS

<222> (73)..(1143)

<223>

<400> 83

gggaaggcga gcagtgccaa tctacagcga agaaagtctc gtttggtaaa agcgagaggg 60

gaaagcctga gc atg cag agt gtg cag agc acg agc ttt tgt ctc cga aag 111  
 Met Gln Ser Val Gln Ser Thr Ser Phe Cys Leu Arg Lys  
 1 5 10

cag tgc ctt tgc ctg acc ttc ctg ctt ctc cat ctc ctg gga cag gtc 159  
 Gln Cys Leu Cys Leu Thr Phe Leu Leu Leu His Leu Leu Gly Gln Val  
 15 20 25

gct gcg act cag cgc tgc cct ccc cag tgc ccg ggc cgg tgc cct gcg 207

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|--|
| Ala | Ala | Thr | Gln | Arg | Cys | Pro | Pro | Gln | Cys | Pro | Gly | Arg | Cys | Pro | Ala |      |  |
| 30  |     |     |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |      |  |
| acg | ccg | ccg | acc | tgc | gcc | ccc | ggg | gtg | cgc | gcg | gtg | ctg | gac | ggc | tgc | 255  |  |
| Thr | Pro | Pro | Thr | Cys | Ala | Pro | Gly | Val | Arg | Ala | Val | Leu | Asp | Gly | Cys |      |  |
|     |     |     |     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |      |  |
| tca | tgc | tgt | ctg | gtg | tgt | gcc | cgc | cag | cgt | ggc | gag | agc | tgc | tca | gat | 303  |  |
| Ser | Cys | Cys | Leu | Val | Cys | Ala | Arg | Gln | Arg | Gly | Glu | Ser | Cys | Ser | Asp |      |  |
|     |     |     |     | 65  |     |     |     | 70  |     |     |     |     | 75  |     |     |      |  |
| ctg | gag | cca | tgc | gac | gag | agc | agt | ggc | ctc | tac | tgt | gat | cgc | agc | gcg | 351  |  |
| Leu | Glu | Pro | Cys | Asp | Glu | Ser | Ser | Gly | Leu | Tyr | Cys | Asp | Arg | Ser | Ala |      |  |
|     |     | 80  |     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |      |  |
| gac | ccc | agc | aac | cag | act | ggc | atc | tgc | acg | gcg | gta | gag | gga | gat | aac | 399  |  |
| Asp | Pro | Ser | Asn | Gln | Thr | Gly | Ile | Cys | Thr | Ala | Val | Glu | Gly | Asp | Asn |      |  |
|     | 95  |     |     |     |     | 100 |     |     |     |     | 105 |     |     |     |     |      |  |
| tgt | gtg | ttc | gat | ggg | gtc | atc | tac | cgc | agt | gga | gag | aaa | ttt | cag | cca | 447  |  |
| Cys | Val | Phe | Asp | Gly | Val | Ile | Tyr | Arg | Ser | Gly | Glu | Lys | Phe | Gln | Pro |      |  |
| 110 |     |     |     |     | 115 |     |     |     |     | 120 |     |     |     | 125 |     |      |  |
| agc | tgc | aaa | ttc | cag | tgc | acc | tgc | aga | gat | ggg | cag | att | ggc | tgt | gtg | 495  |  |
| Ser | Cys | Lys | Phe | Gln | Cys | Thr | Cys | Arg | Asp | Gly | Gln | Ile | Gly | Cys | Val |      |  |
|     |     |     |     | 130 |     |     |     |     | 135 |     |     |     |     | 140 |     |      |  |
| ccc | cgc | tgt | cag | ctg | gat | gtg | cta | ctg | cct | gag | cct | aac | tgc | cca | gct | 543  |  |
| Pro | Arg | Cys | Gln | Leu | Asp | Val | Leu | Leu | Pro | Glu | Pro | Asn | Cys | Pro | Ala |      |  |
|     |     |     |     | 145 |     |     |     | 150 |     |     |     |     | 155 |     |     |      |  |
| cca | aga | aaa | gtt | gag | gtg | cct | gga | gag | tgc | tgt | gaa | aag | tgg | atc | tgt | 591  |  |
| Pro | Arg | Lys | Val | Glu | Val | Pro | Gly | Glu | Cys | Cys | Glu | Lys | Trp | Ile | Cys |      |  |
|     |     | 160 |     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |      |  |
| ggc | cca | gat | gag | gag | gat | tca | ctg | gga | ggc | ctt | acc | ctt | gca | gct | tac | 639  |  |
| Gly | Pro | Asp | Glu | Glu | Asp | Ser | Leu | Gly | Gly | Leu | Thr | Leu | Ala | Ala | Tyr |      |  |
|     |     | 175 |     |     |     | 180 |     |     |     |     | 185 |     |     |     |     |      |  |
| agg | cca | gaa | gcc | acc | cta | gga | gta | gaa | gtc | tct | gac | tca | agt | gtc | aac | 687  |  |
| Arg | Pro | Glu | Ala | Thr | Leu | Gly | Val | Glu | Val | Ser | Asp | Ser | Ser | Val | Asn |      |  |
| 190 |     |     |     |     | 195 |     |     |     |     | 200 |     |     |     | 205 |     |      |  |
| tgc | att | gaa | cag | acc | aca | gag | tgg | aca | gca | tgc | tcc | aag | agc | tgt | ggc | 735  |  |
| Cys | Ile | Glu | Gln | Thr | Thr | Glu | Trp | Thr | Ala | Cys | Ser | Lys | Ser | Cys | Gly |      |  |
|     |     |     |     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |      |  |
| atg | ggg | ttc | tcc | acc | cgg | gtc | acc | aat | agg | aac | cgt | caa | tgt | gag | atg | 783  |  |
| Met | Gly | Phe | Ser | Thr | Arg | Val | Thr | Asn | Arg | Asn | Arg | Gln | Cys | Glu | Met |      |  |
|     |     |     |     | 225 |     |     |     | 230 |     |     |     |     | 235 |     |     |      |  |
| ctg | aaa | cag | act | cgg | ctc | tgc | atg | gtg | cgg | ccc | tgt | gaa | caa | gag | cca | 831  |  |
| Leu | Lys | Gln | Thr | Arg | Leu | Cys | Met | Val | Arg | Pro | Cys | Glu | Gln | Glu | Pro |      |  |
|     |     | 240 |     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |      |  |
| gag | cag | cca | aca | gat | aag | aaa | gga | aaa | aag | tgt | ctc | cgc | acc | aag | aag | 879  |  |
| Glu | Gln | Pro | Thr | Asp | Lys | Lys | Gly | Lys | Lys | Cys | Leu | Arg | Thr | Lys | Lys |      |  |
|     |     | 255 |     |     |     | 260 |     |     |     |     | 265 |     |     |     |     |      |  |
| tca | ctc | aaa | gcc | atc | cac | ctg | cag | ttc | aag | aac | tgc | acc | agc | ctg | cac | 927  |  |
| Ser | Leu | Lys | Ala | Ile | His | Leu | Gln | Phe | Lys | Asn | Cys | Thr | Ser | Leu | His |      |  |
| 270 |     |     |     |     | 275 |     |     |     |     | 280 |     |     |     | 285 |     |      |  |
| acc | tac | aag | ccc | agg | ttc | tgt | ggg | gtc | tgc | agt | gat | ggc | cgc | tgc | tgc | 975  |  |
| Thr | Tyr | Lys | Pro | Arg | Phe | Cys | Gly | Val | Cys | Ser | Asp | Gly | Arg | Cys | Cys |      |  |
|     |     |     |     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |      |  |
| act | ccc | cac | aat | acc | aaa | acc | atc | cag | gca | gag | ttt | cag | tgc | tcc | cca | 1023 |  |

| Thr        | Pro        | His         | Asn        | Thr        | Lys         | Thr        | Ile | Gln        | Ala        | Glu        | Phe | Gln | Cys | Ser | Pro |      |
|------------|------------|-------------|------------|------------|-------------|------------|-----|------------|------------|------------|-----|-----|-----|-----|-----|------|
| 305        |            |             |            |            |             |            |     | 310        |            |            |     |     | 315 |     |     |      |
| ggg        | caa        | ata         | gtc        | aag        | aag         | cca        | gtg | atg        | gtc        | att        | ggg | acc | tgc | acc | tgt | 1071 |
| Gly        | Gln        | Ile         | Val        | Lys        | Lys         | Pro        | Val | Met        | Val        | Ile        | Gly | Thr | Cys | Thr | Cys |      |
|            | 320        |             |            |            |             |            | 325 |            |            |            |     | 330 |     |     |     |      |
| cac        | acc        | aac         | tgt        | cct        | aag         | aac        | aat | gag        | gcc        | ttc        | ctc | cag | gag | ctg | gag | 1119 |
| His        | Thr        | Asn         | Cys        | Pro        | Lys         | Asn        | Asn | Glu        | Ala        | Phe        | Leu | Gln | Glu | Leu | Glu |      |
|            | 335        |             |            |            |             | 340        |     |            |            |            | 345 |     |     |     |     |      |
| ctg        | aag        | act         | acc        | aga        | ggg         | aaa        | atg | taacctatca | ctcaagaagc | acacctacag |     |     |     |     |     | 1173 |
| Leu        | Lys        | Thr         | Thr        | Arg        | Gly         | Lys        | Met |            |            |            |     |     |     |     |     |      |
| 350        |            |             |            |            | 355         |            |     |            |            |            |     |     |     |     |     |      |
| agcacctgta | gctgctgcgc | caccacccat  | caaaggaata | taagaaaagt | aatgaagaat  |            |     |            |            |            |     |     |     |     |     | 1233 |
| cacgatttca | tccttgaatc | ctatgtat    | ttt        | tcctaatgtg | atcatatgag  | gacctttcat |     |            |            |            |     |     |     |     |     | 1293 |
| atctgtcttt | tatttaacaa | aaaatgtaat  | taactgtaaa | cttggaatca | aggtaagctc  |            |     |            |            |            |     |     |     |     |     | 1353 |
| aggatatggc | ttaggaatga | cttactttcc  | tgtggtttta | ttacaaatgc | aaatttctat  |            |     |            |            |            |     |     |     |     |     | 1413 |
| aaatttaaga | aaacaagtat | ataatttact  | ttgtagactg | tttcacattg | cactcatcat  |            |     |            |            |            |     |     |     |     |     | 1473 |
| atthtgttgt | gcactagtgc | aattccaaga  | aaatatcact | gtaatgagtc | agtgaagtct  |            |     |            |            |            |     |     |     |     |     | 1533 |
| agaatcatac | ttacatttcc | attgtacaag  | tattacaacc | atatattgag | gttcattggg  |            |     |            |            |            |     |     |     |     |     | 1593 |
| aagattctct | attggtctcc | ttttgggta   | aaccagctct | gaatttccaa | gtcctcaaatc |            |     |            |            |            |     |     |     |     |     | 1653 |
| caaggaaaca | tgcagctctt | caacatgaca  | tccagagatg | actattactt | ttctgtttag  |            |     |            |            |            |     |     |     |     |     | 1713 |
| ttttacacta | ggaaacgtgt | tgtatctaca  | gtaatgaaat | gtttactaag | tggactgggtg |            |     |            |            |            |     |     |     |     |     | 1773 |
| tcataaactt | tctccattta | agacacattg  | actcctttcc | aatagaaaga | aactaaacag  |            |     |            |            |            |     |     |     |     |     | 1833 |
| aaaactccca | atacaaagat | gactgggtccc | tcatagccct | cagacattta | tatatggaa   |            |     |            |            |            |     |     |     |     |     | 1893 |
| gctgctgagg | cccccaagtt | ttttaattaa  | gcagaaacag | catattagca | gggattctct  |            |     |            |            |            |     |     |     |     |     | 1953 |
| catctaactg | atgagtaaac | tgaggcccaa  | agcacttgct | tacatctct  | gatagctgtt  |            |     |            |            |            |     |     |     |     |     | 2013 |
| tcaaatgtgc | atthtgtgga | atthtgagaa  | aaatagagca | aatcaacat  | gactgggtgtt |            |     |            |            |            |     |     |     |     |     | 2073 |
| gagagaccac | acattttatg | agagtthtga  | attattgtag | acatgcccaa | aacttatcct  |            |     |            |            |            |     |     |     |     |     | 2133 |
| tgggccataa | ttatgaaaac | tcattgatcaa | gatatatgtg | tatacataca | tgtatctggt  |            |     |            |            |            |     |     |     |     |     | 2193 |
| ttgtcaggct | acaaggtagg | ctgcaaaatt  | aatcttagac | attcttttaa | tgccaccaca  |            |     |            |            |            |     |     |     |     |     | 2253 |
| cgtgttccgc | ttctctcttt | taaagtatth  | ataaaaatat | aaattgtaca | ttttgtaaaa  |            |     |            |            |            |     |     |     |     |     | 2313 |
| tattatgttt | gatttctcta | cttgtcatat  | cactaaataa | acacgatttt | attgttgaaa  |            |     |            |            |            |     |     |     |     |     | 2373 |
| aaaaaaaaaa | aaaaaa     |             |            |            |             |            |     |            |            |            |     |     |     |     |     | 2389 |

&lt;210&gt; 84

&lt;211&gt; 357

&lt;212&gt; PRT

&lt;213&gt; NM\_002514 NOV1, nephroblastoma overexpressed gene

&lt;400&gt; 84

Met Gln Ser Val Gln Ser Thr Ser Phe Cys Leu Arg Lys Gln Cys Leu  
1 5 10 15

Cys Leu Thr Phe Leu Leu Leu His Leu Leu Gly Gln Val Ala Ala Thr  
20 25 30

Gln Arg Cys Pro Pro Gln Cys Pro Gly Arg Cys Pro Ala Thr Pro Pro  
35 40 45

Thr Cys Ala Pro Gly Val Arg Ala Val Leu Asp Gly Cys Ser Cys Cys  
50 55 60

Leu Val Cys Ala Arg Gln Arg Gly Glu Ser Cys Ser Asp Leu Glu Pro  
65 70 75 80

Cys Asp Glu Ser Ser Gly Leu Tyr Cys Asp Arg Ser Ala Asp Pro Ser  
85 90 95

Asn Gln Thr Gly Ile Cys Thr Ala Val Glu Gly Asp Asn Cys Val Phe  
100 105 110

Asp Gly Val Ile Tyr Arg Ser Gly Glu Lys Phe Gln Pro Ser Cys Lys  
115 120 125

Phe Gln Cys Thr Cys Arg Asp Gly Gln Ile Gly Cys Val Pro Arg Cys  
130 135 140

Gln Leu Asp Val Leu Leu Pro Glu Pro Asn Cys Pro Ala Pro Arg Lys  
145 150 155 160

Val Glu Val Pro Gly Glu Cys Cys Glu Lys Trp Ile Cys Gly Pro Asp  
165 170 175

Glu Glu Asp Ser Leu Gly Gly Leu Thr Leu Ala Ala Tyr Arg Pro Glu  
180 185 190

Ala Thr Leu Gly Val Glu Val Ser Asp Ser Ser Val Asn Cys Ile Glu  
195 200 205

Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser Cys Gly Met Gly Phe  
210 215 220

Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys Glu Met Leu Lys Gln  
225 230 235 240

Thr Arg Leu Cys Met Val Arg Pro Cys Glu Gln Glu Pro Glu Gln Pro  
245 250 255

Thr Asp Lys Lys Gly Lys Lys Cys Leu Arg Thr Lys Lys Ser Leu Lys  
260 265 270

Ala Ile His Leu Gln Phe Lys Asn Cys Thr Ser Leu His Thr Tyr Lys  
275 280 285

Pro Arg Phe Cys Gly Val Cys Ser Asp Gly Arg Cys Cys Thr Pro His  
290 295 300

Asn Thr Lys Thr Ile Gln Ala Glu Phe Gln Cys Ser Pro Gly Gln Ile  
305 310 315 320

Val Lys Lys Pro Val Met Val Ile Gly Thr Cys Thr Cys His Thr Asn  
325 330 335

Cys Pro Lys Asn Asn Glu Ala Phe Leu Gln Glu Leu Glu Leu Lys Thr  
340 345 350

Thr Arg Gly Lys Met  
355

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU03/01166

**A. CLASSIFICATION OF SUBJECT MATTER**Int. Cl. <sup>7</sup>: C12Q 1/68, G01N 33/53

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

SEE BELOW

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SEE BELOW

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

DGENE, EMBL, GENBANK: SEQ ID NOS 1-7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 46, 48, 50, 52, 53, 55, 57, 59, 60-63, 65, 67, 69, 71-75, 77, 79, 81, 83 WPIDS: ovarian, cancer, tumour, detect, diagnose

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No.                                    |
|-----------|--|--|
| X         | Weitzel JN et al (1994) Gynecologic Oncology 55, pages 245-252 "Molecular genetic changes associated with ovarian cancer" and GENBANK ACC M62397<br>See table 3 and pages 249 and 50 with respect to SEQ ID NOS 3, 4 (MCC) | 1-5, 7, 9-12,<br>16-20, 22, 24-<br>27, 31-43, 62-<br>69  |
| X         | DGENE Abstract Accession No ABQ54626 & WO 02 00677 A1 (HUMAN GENOME SCIENCES, INC.) 3 January 2002<br>See abstract and claim 1, SEQ ID NO 506 (applicant's SEQ ID NO 9)  | 1-5, 7, 10-12,<br>16-20, 22, 25-<br>27, 31-43, 60-<br>69 |

☒ Further documents are listed in the continuation of Box C☒ See patent family annex

|   |  |
|---|--|
| * Special categories of cited documents:  |  |
| "A" document defining the general state of the art which is not considered to be of particular relevance  | "T" later document published after the international filing date, or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention   |
| "E" earlier application or patent but published on or after the international filing date   | "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone   |
| "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| "O" document referring to an oral disclosure, use, exhibition or other means  | "&" document member of the same patent family  |
| "P" document published prior to the international filing date but later than the priority date claimed  |  |

Date of the actual completion of the international search  
23 October 2003Date of mailing of the international search report  
06 NOV 2003Name and mailing address of the ISA/AU  
AUSTRALIAN PATENT OFFICE  
PO BOX 200, WODEN ACT 2606, AUSTRALIA  
E-mail address: pct@ipaustalia.gov.au  
Facsimile No. (02) 6285 3929Authorized officer  
**TERRY MOORE**  
Telephone No : (02) 6283 2632

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU03/01166

| C (Continuation 1). DOCUMENTS CONSIDERED TO BE RELEVANT |  |  |
|---|--|--|
| Category*   | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No.  |
| X   | Hough CD et al (2000) Cancer Research 60, 6281-7 "Large-scale serial analysis of gene expression reveals genes differentially expressed in ovarian cancer"<br>See table 3: mal2, claudin 3(relevant to SEQ ID NOs 11, 15)  | 1-5, 7, 10-12,<br>16-20, 22, 25-<br>27, 31-43, 60-<br>69           |
| PX  | WO 02 071928 A2 (MILLENNIUM PHARMACEUTICALS, INC) 19 September 2002<br>See table 1 pg 20 KIAA0869: pg 19 FLJ22252: pg 21 SLP1, pgs 21, 24, 25 PAX8: pg 18 ANXA2 (applicant's SEQ ID NOs 13, 17, 25, 43 and 71/72) & DGENE Accession No ABS76418 (applicant's SEQ ID NO 17)   | 1-5, 7, 10-12,<br>16-20, 22, 25-<br>27, 31-43, 45,<br>47-53, 60-69 |
| PX  | WO 02 102235 A2 (EOS BIOTECHNOLOGY, INC) 27 December 2002<br>See pg 187 KIAA0869: pg 184 claudin 7: pgs 111, 162 and 180 KIAA0101: pg 120 KIAA1481, pgs 152, 174, 184, 202 paired box gene 8: pgs 155, 161 methylene tetrahydrofolate dehydrogenase: pg 176 DD5: pg 125 SOCS3: pg 261 serum amyloid A1 (applicant's SEQ ID NOs 13, 21, 23, 37, 43, 55, 63, 73/74 and 75) | 1-5, 7, 10-12,<br>16-20, 22, 25-<br>27, 31-43, 45,<br>47-53, 60-69 |
| PX  | DGENE Abstract Accession No ACA66405 & US 2003/0004102 A1 (ASHKENAZI AJ et al) 2 January 2003<br>See abstract and claim 2, figure 203 (applicant's SEQ ID NO 19)   | 1-5, 7, 10-12,<br>16-20, 22, 25-<br>27, 31-43, 60-<br>69           |
| X   | Scheurle D et al "Cancer gene discovery using digital differential display" (2000)(online), (retrieved on 16 October 2003) Retrieved from Internet<URL:http://www.fau.edu/cmhb/publications/cancergenes.htm><br>See "up-regulation of known genes: claudin 7, secreted frizzled rel. pro. 4 (applicant's SEQ ID NOs 21 and 69)   | 1-5, 7, 10-12,<br>16-20, 22, 25-<br>27, 31-43, 45,<br>47-53, 60-69 |
| X   | DGENE Abstract Accession No AAS56533 & WO 01 70796 A2 (CORIXA CORPORATION) 27 September 2001<br>See abstract and claim 1, page 165 KIAA0101 (applicant's SEQ ID NO 23)   | 1-5, 7, 10-12,<br>16-20, 22, 25-<br>27, 31-43, 60-<br>69           |
| X   | Medline Abstract 11906548 Shigemasa K et al (2001) Int J Gynecol Cancer 11(6), 454-61 "Expression of the protease inhibitor antileukoprotease and the serine protease stratum corneum chymotryptic enzyme (SCCE) is coordinated in ovarian tumours"<br>See abstract, relevant to applicant's SEQ ID NO 25  | 1-5, 7, 9-11,<br>16-20, 22, 25-<br>27, 31-42, 62-<br>69            |



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU03/01166

| C (Continuation 2) DOCUMENTS CONSIDERED TO BE RELEVANT |  |  |
|--|--|--|
| Category*  | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No.  |
| X  | Laval S et al (1994) Cell Growth Differ 5(11), 1173-83 "Isolation and characterization of an epithelial-specific receptor tyrosine kinase from an ovarian cancer cell line"<br>See whole document, relevant to applicant's SEQ ID NO 27  | 1-5, 7, 9-11, 16-20, 22, 25-27, 31-42, 62-69                 |
| PX   | WO 03 068054 A2 (THE GOVERNMENT OF THE UNITED STATES OF AMERICA as represented by THE SECRETARY, DEPARTMENT OF HEALTH SERVICES) 21 August 2003<br>See in particular table 1, pg 22/60 EDDR1: pg 32/60 TOP2A: 29/60 PAX8 and 33/60 SFRP4 (applicant's SEQ ID NOs 27, 29, 43 and 69)   | 1-5, 7, 10-12, 16-20, 22, 25-27, 31-43, 45, 47, 51-53, 62-69 |
| X  | Koshiyama M et al (2001) Anticancer Res 21, 2925-32 "Immunohistochemical expression of topoisomerase II alpha (TopII alpha) and multidrug resistance-associated protein (MRP), plus chemosensitivity testing, as chemotherapeutic indices of ovarian endometrial carcinomas"<br>See whole document, relevant to applicant's SEQ ID NO 29 | 1-5, 7, 10-12, 16-20, 22, 26, 27, 31-40, 42, 43, 62-69       |
| X  | Gotlieb WH et al (2001) Gynecol Oncol 82(1), 99-104 "Topoisomerase II immunostaining as a prognostic marker for survival in ovarian cancer"<br>See whole document, relevant to applicant's SEQ ID NO 29  | 1-5, 7, 10-12, 16-20, 22, 26, 27, 31-40, 42, 43, 62-69       |
| X  | Costa MJ et al (2000) Int J Gynecol Pathol 19(3), 248-57 "Topoisomerase II alpha: prognostic predictor and cell cycle marker in surface epithelial neoplasms of the ovary and peritoneum"<br>See whole document, relevant to applicant's SEQ ID NO 29  | 1-5, 7, 10-12, 16-20, 22, 25-27, 31-43, 62-69                |
| X  | DGENE Abstract Accession No ABQ54317 & WO 02 00677 A1 (HUMAN GENOME SCIENCES, INC.) 3 January 2002<br>See abstract and claim 1, SEQ ID NO 197 (applicant's SEQ ID NO 31 reverse complement)  | 1-5, 7, 10-12, 16-20, 22, 25-27, 31-43, 60-69                |
| PX   | DGENE Abstract Accession No ACA03948 & US 2002/0192678 A1 (CHEN H-M) 19 December 2002<br>See abstract and example 13, pages 98-99 (applicant's SEQ ID NO 33)   | 1-5, 7, 10-12, 16-20, 22, 24-27, 31-43, 62-69                |

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU03/01166

| C (Continuation 3) DOCUMENTS CONSIDERED TO BE RELEVANT |  |  |
|--|--|--|
| Category*  | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No.                                    |
| X  | Medline Abstract 11925933 Sakamoto M et al (2001) Hum Cell 14(4), 305-15<br>"Analysis of gene expression profiles associated with cisplatin resistance in human ovarian cancer cell lines and tissues using cDNA microarray"<br>See whole document, relevant to applicant's SEQ ID NO 35 | 1-5, 7, 10-12,<br>16-20, 22, 24-<br>27, 31-43, 62-<br>69 |
| X  | DGENE Abstract Accession No ABQ54876 & WO 02 00677 A1 (HUMAN GENOME SCIENCES, INC.) 3 January 2002<br>See abstract and claim 1, SEQ ID NO 756 (applicant's SEQ ID NO 39)   | 1-5, 7, 10-12,<br>16-20, 22, 25-<br>27, 31-43, 60-<br>69 |
| X  | DGENE Abstract Accession No ABL83226 & WO 01 92581 A2 (CORIXA CORPORATION) 6 December 2001<br>See abstract and claim 1, SEQ ID NO 6204 (applicant's SEQ ID NO 41)  | 1-5, 7, 10-12,<br>16-20, 22, 25-<br>27, 31-43, 60-<br>69 |
| X  | Genbank Accession No AK075046 3 September 2002<br>Isogai T et al <i>Homo sapiens</i> cDNA FLJ90565, clone OVARC1001336<br>See "features" (applicant's SEQ ID NO 50)  | 1-5, 7, 16-20,<br>22, 31-40, 62-<br>67                   |
| X  | DGENE Abstract Accession No AAH83197 & WO 01 51513 A2 (CORIXA CORPORATION) 19 July 2001<br>See abstract and claim 5, SEQ ID NO 821 (applicant's SEQ ID NO 52)  | 1-5, 7, 10-12,<br>16-20, 22, 25-<br>27, 31-43, 60-<br>69 |
| X  | DGENE Abstract Accession No ABL90699 & WO 01 90304 A2 (HUMAN GENOME SCIENCES) 29 November 2001<br>See abstract and claim 4, SEQ ID NO 1261 (applicant's SEQ ID NO 53)  | 1-5, 7, 10-12,<br>16-20, 22, 25-<br>27, 31-43, 60-<br>69 |
| X  | DGENE Abstract Accession No AAZ77553 & WO 99 53040 A2 (METAGEN GESELLSCHAFT F R GENOMFORSCHUNG) 21 October 1999<br>See abstract, claim 3 and page 226 (applicant's SEQ ID NO 57)   | 1-5, 7, 10-12,<br>15-20, 22, 25-<br>27, 30-43, 60-<br>69 |
| X  | DGENE Abstract Accession No ABL67667 & WO 01 94629 A2 (AVALON PHARMACEUTICALS) 13 December 2001<br>See abstract and claim 1, SEQ ID NO 6004 (applicant's SEQ ID NOs 59, 60)  | 1-5, 7, 10-12,<br>15-20, 22, 25-<br>27, 30-43, 60-<br>69 |

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU03/01166

| C (Continuation 4) DOCUMENTS CONSIDERED TO BE RELEVANT |   |  |
|--|---|--|
| Category*  | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No.  |
| X  | DGENE Abstract Accession No ABL61830 & WO 01 94629 A2 (AVALON PHARMACEUTICALS) 13 December 2001<br>See abstract and claim 1, SEQ ID NO 167 (applicant's SEQ ID NOs 61, 62)                                      | 1-5, 7, 10-12, 15-20, 22, 24-27, 30-43, 62-69                |
| X  | DGENE Abstract Accession No AAV64871 & WO 98 48010 A1 (GARVAN INSTITUTE OF MEDICAL RESEARCH) 29 October 1998<br>See Abstract and figure 3B, example and pg 22 line 34 of WO 98 48010 (applicant's SEQ ID NO 63) | 1-5, 7, 10-12, 16-20, 22, 25-27, 31-43, 45, 47, 51-53, 60-69 |
| X  | DGENE Abstract Accession No ABV34360 & WO 01 60860 A2 (MILLENNIUM PREDICTIVE MEDICINE, INC) 23 August 2001<br>See Abstract, claim 1, pg 19 and pg 7222 (applicant's SEQ ID NO 65)                               | 1-5, 7, 10-12, 16-20, 22, 25-27, 31-43, 45, 47, 51-53, 60-69 |
| X  | DGENE Abstract Accession No ABQ54583 & WO 02 00677 A1 (HUMAN GENOME SCIENCES, INC.) 3 January 2002<br>See in particular abstract and claim 1, SEQ ID NO 463 (applicant's SEQ ID NO 67)                          | 1-5, 7, 10-12, 16-20, 22, 25-27, 31-43, 45, 47, 51-53, 60-69 |
| X  | DGENE Abstract Accession No ABL62031 & WO 01 94629 A2 (AVALON PHARMACEUTICALS) 13 December 2001<br>See abstract and claim 1, SEQ ID NO 368 (applicant's SEQ ID NOs 73, 74)                                      | 1-5, 7, 10-12, 16-20, 22, 25-27, 31-43, 45, 47-53, 60-69     |
| X  | DGENE Abstract Accession No AAI99253 & WO 01 55313 A2 (HUMAN GENOME SCIENCES, INC) 2 August 2001<br>See abstract, example 2 and SEQ ID NO 1017 (applicant's SEQ ID NO 77)                                       | 1-5, 7, 10-12, 16-20, 22, 25-27, 31-43, 45, 47, 51-53, 60-69 |
| X  | DGENE Abstract Accession No ABL64081 & WO 01 94629 A2 (AVALON PHARMACEUTICALS) 13 December 2001<br>See abstract and claim 1, SEQ ID NO 2418 (applicant's SEQ ID NO 79)  | 1-5, 7, 10-12, 16-20, 22, 25-27, 31-43, 45, 47-53, 60-69     |

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/AU03/01166

| C (Continuation 5) DOCUMENTS CONSIDERED TO BE RELEVANT |  |  |
|--|--|--|
| Category*  | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No.  |
| X  | DGENE Abstract Accession No ABL66499 & WO 01 94629 A2 (AVALON PHARMACEUTICALS) 13 December 2001<br>See abstract and claim 1, SEQ ID NO 4836 (applicant's SEQ ID NO 81) | 1-5, 7, 10-12,<br>16-20, 22, 24-<br>27, 31-43, 45,<br>47, 51-53, 60-<br>69 |

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/AU03/01166

## Box I Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos :  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos : 70-73  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:  
The claims are not limited to the technical features of the invention as disclosed in, and supported by, the specification. The claims relate to methods of assessing the promoter regions of specific genes with respect to hypermethylation and ovarian cancer. However the specification only discloses sequences derived from cDNA clones and does not disclose any promoter sequences.
3. ☐ Claims Nos :  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

## Box II Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See supplementary sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.: 2-5, 7, 9, 15, 19, 20, 22, 24, 30, 45 and 47-50 in full and claims 1, 10-12, 16-18, 25-27, 31-43, 51-53 and 60-69 as they relate to table 3.
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

### Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☒ No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU03/01166

**Supplemental Box**

(To be used when the space in any of Boxes I to VIII is not sufficient)

**Continuation of Box No: II**

The application claims more than one invention. Rule 13.1 of the PCT states the principle that an International Application should relate to only one invention or, if there is more than one invention, that the inclusion of those inventions in one International Application is only permitted if all inventions are so linked as to form a single general inventive concept.

Rule 13.2 of the PCT defines the method for determining whether the requirement of unity of invention is satisfied in respect of a group of inventions claimed in an International application. Unity of invention exists only when there is a technical relationship among the claimed inventions involving one or more of the same or corresponding "special technical features." The expression "special technical features" is defined in Rule 13.2 as meaning those technical features that define a contribution which each of the inventions, considered as a whole, makes over the prior art. The determination is made on the contents of the claims as interpreted in light of the description and drawings (if any).

The claims and the description relate to methods for the diagnosis, prediction and treatment of ovarian cancer, using gene and peptide sequences whose expression is altered in ovarian cancer. In particular the claims define methods relating to the use of over 693 specific gene and/or peptide sequences.

Although all of the claims share the common feature that they all relate to methods or apparatus that involve the use of specific gene or peptide sequences whose expression is altered in ovarian cancer, this feature is known (see the documents listed below). As such this feature cannot be regarded as a special technical feature and cannot confer unity on the inventions relating to the use of specific gene or peptide sequences for the detection or treatment of ovarian cancer.

D1 US 5 912 142

D2 US 6 268 165

D3 WO 01 21653

D4 Bayani et al (2002) Cancer Research 62, 3466-76

Furthermore, there is nothing in the specification to suggest that the genes can be further divided into groups, where sequences within a group share a common special technical feature. Although the specification discloses that some of the genes can be classified into further, narrower groups, these further divisions also cannot be regarded as special technical features. These divisions correspond to groups of genes that are down-regulated or up-regulated in cancers and genes that are associated with specific types of ovarian cancers, such as epithelial or mucinous ovarian cancer. However these groups are known, as are methods of using genes specifically associated with these narrower groups for the treatment and detection of specific subsets of ovarian cancers. See in particular D4

Although there is a lack of unity the ISA has searched, as a service to the applicant, five sets of ten sequences for five search fees.

These sequences comprise SEQ ID NOS 1- 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 46, 48, 50, 52, 53, 55, 57, 59, 60- 63, 65, 67, 69, 71-75, 77, 79, 81, 83.

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
**PCT/AU03/01166**

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

| Patent Document Cited in<br>Search Report |         | Patent Family Member |          |    |          |
|---|---------|----------------------|----------|----|----------|
| WO  | 0200677 | AU                   | 29508/01 | AU | 30958/01 |
|   |         | AU                   | 36460/01 | AU | 36461/01 |
|   |         | AU                   | 36463/01 | AU | 36464/01 |
|   |         | AU                   | 36466/01 | AU | 37943/01 |
|   |         | AU                   | 37947/01 | AU | 37949/01 |
|   |         | AU                   | 37951/01 | AU | 37952/01 |
|   |         | AU                   | 37954/01 | AU | 37955/01 |
|   |         | AU                   | 37958/01 | AU | 38585/01 |
|   |         | AU                   | 39727/01 | AU | 39728/01 |
|   |         | AU                   | 41403/01 | AU | 41404/01 |
|   |         | AU                   | 41406/01 | AU | 41407/01 |
|   |         | AU                   | 41409/01 | AU | 41410/01 |
|   |         | AU                   | 41412/01 | AU | 41413/01 |
|   |         | AU                   | 41415/01 | AU | 41416/01 |
|   |         | AU                   | 41418/01 | AU | 41419/01 |
|   |         | AU                   | 43135/01 | AU | 43136/01 |
|   |         | AU                   | 45262/01 | AU | 45354/01 |
|   |         | AU                   | 47191/01 | AU | 49052/01 |
|   |         | AU                   | 49054/01 | AU | 50767/01 |
|   |         | AU                   | 50769/01 | AU | 50770/01 |
|   |         | AU                   | 50772/01 | AU | 50771/01 |
|   |         | AU                   | 50772/01 | AU | 52878/01 |
|   |         | AU                   | 55162/01 | AU | 57912/01 |
|   |         | AU                   | 62899/01 | AU | 66787/01 |
|   |         | CA                   | 2392422  | CA | 2392428  |
|   |         | CA                   | 2392450  | CA | 2392751  |
|   |         | CA                   | 2393002  | CA | 2393616  |
|   |         | CA                   | 2393652  | CA | 2393912  |
|   |         | CA                   | 2393954  | CA | 2394022  |
|   |         | CA                   | 2394841  | CA | 2395178  |
|   |         | CA                   | 2395398  | CA | 2395403  |
|   |         | CA                   | 2395666  | CA | 2395671  |
|   |         | CA                   | 2395693  | CA | 2395699  |
|   |         | AU                   | 36459/01 | AU | 36462/01 |
|   |         | AU                   | 36465/01 | AU | 37944/01 |
|   |         | AU                   | 37950/01 | AU | 37953/01 |
|   |         | AU                   | 37957/01 | AU | 39726/01 |
|   |         | AU                   | 41402/01 | AU | 41405/01 |
|   |         | AU                   | 41408/01 | AU | 41411/01 |
|   |         | AU                   | 41414/01 | AU | 41417/01 |
|   |         | AU                   | 43134/01 | AU | 43137/01 |
|   |         | AU                   | 47190/01 | AU | 49053/01 |
|   |         | AU                   | 50768/01 | AU | 50771/01 |
|   |         | AU                   | 52879/01 | AU | 60969/01 |
|   |         | CA                   | 2392398  | CA | 2392438  |
|   |         | CA                   | 2392757  | CA | 2393618  |
|   |         | CA                   | 2393941  | CA | 2394039  |
|   |         | CA                   | 2395295  | CA | 2395654  |
|   |         | CA                   | 2395676  | CA | 2395724  |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/AU03/01166

|    |            |    |            |    |            |
|----|------------|----|------------|----|------------|
| CA | 2395729    | CA | 2395734    | CA | 2395738    |
| CA | 2395787    | CA | 2395794    | CA | 2395811    |
| CA | 2395815    | CA | 2395816    | CA | 2395827    |
| CA | 2395838    | CA | 2395849    | CA | 2395857    |
| CA | 2395858    | CA | 2395872    | CA | 2395885    |
| CA | 2395889    | CA | 2396719    | CA | 2397407    |
| CA | 2397502    | CA | 2397839    | CA | 2398227    |
| CA | 2398275    | CA | 2398411    | CA | 2398877    |
| EP | 1251863    | EP | 1252176    | EP | 1252185    |
| EP | 1252289    | EP | 1252290    | EP | 1252297    |
| EP | 1252302    | EP | 1252303    | EP | 1252312    |
| EP | 1252326    | EP | 1252337    | EP | 1254147    |
| EP | 1254148    | EP | 1254150    | EP | 1254151    |
| EP | 1254152    | EP | 1254153    | EP | 1254154    |
| EP | 1254157    | EP | 1254165    | EP | 1254170    |
| EP | 1254171    | EP | 1254172    | EP | 1254173    |
| EP | 1254217    | EP | 1254218    | EP | 1254219    |
| EP | 1254228    | EP | 1254242    | EP | 1254248    |
| EP | 1254272    | EP | 1255766    | EP | 1255767    |
| EP | 1255768    | EP | 1255776    | EP | 1255777    |
| EP | 1255778    | EP | 1255817    | EP | 1255864    |
| EP | 1255869    | EP | 1259525    | EP | 1259526    |
| EP | 1259528    | EP | 1259531    | EP | 1259540    |
| EP | 1259642    | EP | 1261380    | EP | 1261618    |
| EP | 1261619    | EP | 1261633    | EP | 1261634    |
| EP | 1261635    | EP | 1261637    | EP | 1261703    |
| EP | 1261742    | EP | 1261745    | EP | 1263944    |
| EP | 1265910    | EP | 1268527    | US | 2002042096 |
| US | 2002042386 | US | 2002061521 | US | 2002077270 |
| US | 2002086330 | US | 2002086353 | US | 2002086811 |
| US | 2002086820 | US | 2002086821 | US | 2002086822 |
| US | 2002086823 | US | 2002090615 | US | 2002090672 |
| US | 2002090673 | US | 2002090674 | US | 2002094953 |
| US | 2002102638 | US | 2002119919 | US | 2002120103 |
| US | 2002132753 | US | 2002132767 | US | 2002147140 |
| US | 2002151009 | US | 2002151479 | US | 2002160493 |
| US | 2002161208 | US | 2002164685 | US | 2002165137 |



**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
**PCT/AU03/01166**

|    |            |    |            |    |            |
|----|------------|----|------------|----|------------|
| US | 2002168711 | US | 2002173454 | US | 2003013649 |
| US | 2003039993 | US | 2003039994 | US | 2003044890 |
| US | 2003044904 | US | 2003044905 | US | 2003044907 |
| US | 2003049650 | US | 2003049652 | US | 2003049703 |
| US | 2003050231 | US | 2003054368 | US | 2003054373 |
| US | 2003054375 | US | 2003054377 | US | 2003054379 |
| US | 2003054420 | US | 2003059875 | US | 2003059908 |
| US | 2003077602 | US | 2003077606 | US | 2003077703 |
| US | 2003077704 | US | 2003077808 | US | 2003082681 |
| US | 2003082758 | US | 2003092102 | US | 2003092611 |
| US | 2003092615 | US | 2003096346 | US | 2003108907 |
| US | 2003125246 | US | 2003139327 | US | 2003157508 |
| US | 2003175739 | US | 2003190707 | WO | 0154472    |
| WO | 0154473    | WO | 0154474    | WO | 0154708    |
| WO | 0154733    | WO | 0155162    | WO | 0155163    |
| WO | 0155164    | WO | 0155167    | WO | 0155168    |
| WO | 0155173    | WO | 0155200    | WO | 0155201    |
| WO | 0155202    | WO | 0155203    | WO | 0155204    |
| WO | 0155205    | WO | 0155206    | WO | 0155207    |
| WO | 0155208    | WO | 0155300    | WO | 0155301    |
| WO | 0155302    | WO | 0155303    | WO | 0155304    |
| WO | 0155305    | WO | 0155306    | WO | 0155307    |
| WO | 0155308    | WO | 0155309    | WO | 0155310    |
| WO | 0155311    | WO | 0155312    | WO | 0155313    |
| WO | 0155314    | WO | 0155315    | WO | 0155316    |
| WO | 0155317    | WO | 0155318    | WO | 0155319    |
| WO | 0155320    | WO | 0155321    | WO | 0155322    |
| WO | 0155323    | WO | 0155324    | WO | 0155325    |
| WO | 0155326    | WO | 0155327    | WO | 0155328    |
| WO | 0155329    | WO | 0155343    | WO | 0155350    |
| WO | 0155355    | WO | 0155364    | WO | 0155367    |
| WO | 0155368    | WO | 0155387    | WO | 0155388    |
| WO | 0155430    | WO | 0155440    | WO | 0155441    |
| WO | 0155447    | WO | 0155448    | WO | 0155449    |
| WO | 0157182    | WO | 0159063    | WO | 0159064    |
| WO | 0162789    | WO | 0164703    |    |            |
| WO | 02071928   | US | 2003087250 |    |            |

INTERNATIONAL SEARCH REPORT  
Information on patent family members

International application No.

PCT/AU03/01166

|    |            |    |            |    |            |    |            |
|----|------------|----|------------|----|------------|----|------------|
| WO | 02102235   | AU | 69022/00   | CA | 2381699    | EP | 1204764    |
|    |            | US | 2002015970 | US | 2002019330 | US | 2003124579 |
|    |            | US | 2003152926 | WO | 0111086    | WO | 02079492   |
|    |            | WO | 02086443   | WO | 02098358   | WO | 03003906   |
|    |            | WO | 03025138   | WO | 03042661   |    |            |
| US | 2003004102 | AU | 10703/99   | AU | 15324/99   | AU | 22122/99   |
|    |            | CA | 2305385    | CA | 2309358    | CA | 2321677    |
|    |            | CA | 2347835    | CA | 2353775    | EP | 1025227    |
|    |            | EP | 1032667    | EP | 1135485    | EP | 1141284    |
|    |            | EP | 1141289    | EP | 1159419    | EP | 1159422    |
|    |            | EP | 1173563    | EP | 1220905    | EP | 1263948    |
|    |            | US | 6410708    | US | 2002072067 | US | 2002072092 |
|    |            | US | 2002072496 | US | 2002072497 | US | 2002103125 |
|    |            | US | 2002119118 | US | 2002119130 | US | 2002123463 |
|    |            | US | 2002127576 | US | 2002127584 | US | 2002132252 |
|    |            | US | 2002132253 | US | 2002142961 | US | 2002156006 |
|    |            | US | 2002160384 | US | 2002161199 | US | 2002169284 |
|    |            | US | 2002177164 | US | 2002177165 | US | 2002177188 |
|    |            | US | 2002177553 | US | 2002182638 | US | 2002182673 |
|    |            | US | 2002183493 | US | 2002183494 | US | 2002183505 |
|    |            | US | 2002192209 | US | 2002192706 | US | 2002192752 |
|    |            | US | 2002193299 | US | 2002193300 | US | 2002197615 |
|    |            | US | 2002197674 | US | 2002198148 | US | 2002198149 |
|    |            | US | 2003003531 | US | 2003003546 | US | 2003004311 |
|    |            | US | 2003008297 | US | 2003008352 | US | 2003008353 |
|    |            | US | 2003009012 | US | 2003009013 | US | 2003013855 |
|    |            | US | 2003017476 | US | 2003017540 | US | 2003017541 |
|    |            | US | 2003017542 | US | 2003017543 | US | 2003017544 |
|    |            | US | 2003017563 | US | 2003017981 | US | 2003017982 |
|    |            | US | 2003018168 | US | 2003022187 | US | 2003022239 |
|    |            | US | 2003022293 | US | 2003022294 | US | 2003022295 |
|    |            | US | 2003022296 | US | 2003022297 | US | 2003022298 |
|    |            | US | 2003022300 | US | 2003022301 | US | 2003022328 |
|    |            | US | 2003022331 | US | 2003027162 | US | 2003027163 |
|    |            | US | 2003027212 | US | 2003027262 | US | 2003027263 |
|    |            | US | 2003027264 | US | 2003027265 | US | 2003027266 |
|    |            | US | 2003027267 | US | 2003027269 | US | 2003027270 |

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
**PCT/AU03/01166**

|    |            |    |            |    |            |
|----|------------|----|------------|----|------------|
| US | 2003027271 | US | 2003027272 | US | 2003027273 |
| US | 2003027274 | US | 2003027275 | US | 2003027276 |
| US | 2003027277 | US | 2003027278 | US | 2003027279 |
| US | 2003027280 | US | 2003027281 | US | 2003027324 |
| US | 2003027754 | US | 2003027985 | US | 2003027986 |
| US | 2003027992 | US | 2003027993 | US | 2003032023 |
| US | 2003032057 | US | 2003032062 | US | 2003032063 |
| US | 2003032101 | US | 2003032102 | US | 2003032103 |
| US | 2003032104 | US | 2003032105 | US | 2003032106 |
| US | 2003032107 | US | 2003032108 | US | 2003032109 |
| US | 2003032110 | US | 2003032111 | US | 2003032112 |
| US | 2003032113 | US | 2003032114 | US | 2003032115 |
| US | 2003032116 | US | 2003032117 | US | 2003032118 |
| US | 2003032119 | US | 2003032120 | US | 2003032121 |
| US | 2003032122 | US | 2003032123 | US | 2003032124 |
| US | 2003032125 | US | 2003032126 | US | 2003032127 |
| US | 2003032128 | US | 2003032129 | US | 2003032130 |
| US | 2003032131 | US | 2003032132 | US | 2003032133 |
| US | 2003032134 | US | 2003032135 | US | 2003032136 |
| US | 2003032137 | US | 2003032138 | US | 2003032139 |
| US | 2003032140 | US | 2003032155 | US | 2003032156 |
| US | 2003034993 | US | 2003036117 | US | 2003036118 |
| US | 2003036120 | US | 2003036121 | US | 2003036122 |
| US | 2003036123 | US | 2003036124 | US | 2003036125 |
| US | 2003036126 | US | 2003036127 | US | 2003036128 |
| US | 2003036129 | US | 2003036130 | US | 2003036131 |
| US | 2003036132 | US | 2003036133 | US | 2003036134 |
| US | 2003036135 | US | 2003036136 | US | 2003036137 |
| US | 2003036138 | US | 2003036139 | US | 2003036140 |
| US | 2003036141 | US | 2003036142 | US | 2003036143 |
| US | 2003036144 | US | 2003036145 | US | 2003036146 |
| US | 2003036147 | US | 2003036148 | US | 2003036149 |
| US | 2003036150 | US | 2003036151 | US | 2003036152 |
| US | 2003036153 | US | 2003036154 | US | 2003036155 |
| US | 2003036156 | US | 2003036157 | US | 2003036158 |
| US | 2003036159 | US | 2003036160 | US | 2003036161 |
| US | 2003036162 | US | 2003036164 | US | 2003036165 |

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
**PCT/AU03/01166**

|    |            |    |            |    |            |
|----|------------|----|------------|----|------------|
| US | 2003036179 | US | 2003036180 | US | 2003036634 |
| US | 2003038826 | US | 2003038827 | US | 2003038828 |
| US | 2003039648 | US | 2003040013 | US | 2003040014 |
| US | 2003040054 | US | 2003040055 | US | 2003040056 |
| US | 2003040057 | US | 2003040058 | US | 2003040059 |
| US | 2003040060 | US | 2003040061 | US | 2003040062 |
| US | 2003040063 | US | 2003040064 | US | 2003040065 |
| US | 2003040066 | US | 2003040067 | US | 2003040068 |
| US | 2003040069 | US | 2003040070 | US | 2003040071 |
| US | 2003040072 | US | 2003040073 | US | 2003040074 |
| US | 2003040075 | US | 2003040076 | US | 2003040077 |
| US | 2003040078 | US | 2003040473 | US | 2003043176 |
| US | 2003044806 | US | 2003044841 | US | 2003044844 |
| US | 2003044902 | US | 2003044916 | US | 2003044917 |
| US | 2003044918 | US | 2003044919 | US | 2003044920 |
| US | 2003044921 | US | 2003044922 | US | 2003044923 |
| US | 2003044924 | US | 2003044925 | US | 2003044926 |
| US | 2003044927 | US | 2003044928 | US | 2003044929 |
| US | 2003044930 | US | 2003044931 | US | 2003044932 |
| US | 2003044945 | US | 2003045462 | US | 2003045463 |
| US | 2003045684 | US | 2003049633 | US | 2003049638 |
| US | 2003049681 | US | 2003049682 | US | 2003049684 |
| US | 2003049735 | US | 2003049738 | US | 2003049741 |
| US | 2003049756 | US | 2003049816 | US | 2003049817 |
| US | 2003050239 | US | 2003050240 | US | 2003050241 |
| US | 2003050457 | US | 2003050462 | US | 2003054359 |
| US | 2003054403 | US | 2003054404 | US | 2003054405 |
| US | 2003054406 | US | 2003054442 | US | 2003054454 |
| US | 2003054455 | US | 2003054456 | US | 2003054458 |
| US | 2003054459 | US | 2003054460 | US | 2003054468 |
| US | 2003054516 | US | 2003054517 | US | 2003054986 |
| US | 2003054987 | US | 2003055216 | US | 2003055222 |
| US | 2003059780 | US | 2003059782 | US | 2003059783 |
| US | 2003059831 | US | 2003059832 | US | 2003059833 |
| US | 2003059876 | US | 2003059909 | US | 2003060406 |
| US | 2003060407 | US | 2003060600 | US | 2003060601 |
| US | 2003060602 | US | 2003064062 | US | 2003064375 |

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
**PCT/AU03/01166**

|    |            |    |            |    |            |
|----|------------|----|------------|----|------------|
| US | 2003064407 | US | 2003064440 | US | 2003064441 |
| US | 2003064925 | US | 2003065142 | US | 2003065143 |
| US | 2003068793 | US | 2003068794 | US | 2003068795 |
| US | 2003068796 | US | 2003068797 | US | 2003068798 |
| US | 2003069178 | US | 2003069179 | US | 2003069394 |
| US | 2003069403 | US | 2003071835 | US | 2003072745 |
| US | 2003073090 | US | 2003073129 | US | 2003073130 |
| US | 2003073131 | US | 2003073169 | US | 2003073170 |
| US | 2003073171 | US | 2003073172 | US | 2003073173 |
| US | 2003073174 | US | 2003073177 | US | 2003073208 |
| US | 2003073210 | US | 2003073211 | US | 2003073212 |
| US | 2003073213 | US | 2003073214 | US | 2003073215 |
| US | 2003073216 | US | 2003073624 | US | 2003073809 |
| US | 2003073810 | US | 2003073813 | US | 2003073821 |
| US | 2003077593 | US | 2003077594 | US | 2003077659 |
| US | 2003077700 | US | 2003077718 | US | 2003077732 |
| US | 2003077733 | US | 2003077776 | US | 2003077777 |
| US | 2003077778 | US | 2003077779 | US | 2003077780 |
| US | 2003077781 | US | 2003077782 | US | 2003077783 |
| US | 2003077784 | US | 2003077785 | US | 2003077786 |
| US | 2003077787 | US | 2003077788 | US | 2003077789 |
| US | 2003077790 | US | 2003077791 | US | 2003077792 |
| US | 2003078377 | US | 2003082199 | US | 2003082546 |
| US | 2003082626 | US | 2003082627 | US | 2003082628 |
| US | 2003082689 | US | 2003082715 | US | 2003082716 |
| US | 2003082717 | US | 2003082759 | US | 2003082760 |
| US | 2003082761 | US | 2003082762 | US | 2003082763 |
| US | 2003082764 | US | 2003082765 | US | 2003082766 |
| US | 2003082767 | US | 2003083248 | US | 2003083461 |
| US | 2003083462 | US | 2003087304 | US | 2003087305 |
| US | 2003087344 | US | 2003087345 | US | 2003087365 |
| US | 2003087367 | US | 2003087373 | US | 2003087374 |
| US | 2003092061 | US | 2003092104 | US | 2003092147 |
| US | 2003092883 | US | 2003096353 | US | 2003096386 |
| US | 2003096744 | US | 2003096954 | US | 2003096955 |
| US | 2003099625 | US | 2003100061 | US | 2003100087 |
| US | 2003104536 | US | 2003104538 | US | 2003104539 |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/AU03/01166

|    |            |    |            |    |            |
|----|------------|----|------------|----|------------|
| US | 2003104540 | US | 2003104541 | US | 2003104542 |
| US | 2003104543 | US | 2003104998 | US | 2003109672 |
| US | 2003113793 | US | 2003113795 | US | 2003114652 |
| US | 2003119001 | US | 2003119055 | US | 2003119097 |
| US | 2003119105 | US | 2003119106 | US | 2003120053 |
| US | 2003124531 | US | 2003124661 | US | 2003125535 |
| US | 2003129650 | US | 2003129695 | US | 2003130181 |
| US | 2003130182 | US | 2003130483 | US | 2003130490 |
| US | 2003130491 | US | 2003134284 | US | 2003134785 |
| US | 2003135034 | US | 2003138882 | US | 2003138898 |
| US | 2003139328 | US | 2003139329 | US | 2003148373 |
| US | 2003148375 | US | 2003148376 | US | 2003153000 |
| US | 2003157635 | US | 2003166104 | US | 2003166105 |
| US | 2003166106 | US | 2003166107 | US | 2003166108 |
| US | 2003166109 | US | 2003166110 | US | 2003166111 |
| US | 2003166112 | US | 2003166113 | US | 2003166114 |
| US | 2003166115 | US | 2003166117 | US | 2003166118 |
| US | 2003166119 | US | 2003166120 | US | 2003166121 |
| US | 2003166122 | US | 2003166123 | US | 2003166124 |
| US | 2003166125 | US | 2003166126 | US | 2003166127 |
| US | 2003166128 | US | 2003166129 | US | 2003166130 |
| US | 2003166131 | US | 2003166133 | US | 2003166134 |
| US | 2003166135 | US | 2003166188 | US | 2003166848 |
| US | 2003170718 | US | 2003170721 | US | 2003175882 |
| US | 2003180311 | US | 2003180792 | US | 2003180793 |
| US | 2003180794 | US | 2003180795 | US | 2003180796 |
| US | 2003180836 | US | 2003180837 | US | 2003180838 |
| US | 2003180839 | US | 2003180840 | US | 2003180841 |
| US | 2003180842 | US | 2003180843 | US | 2003180844 |
| US | 2003180845 | US | 2003180846 | US | 2003180847 |
| US | 2003180848 | US | 2003180849 | US | 2003180850 |
| US | 2003180851 | US | 2003180852 | US | 2003180853 |
| US | 2003180854 | US | 2003180855 | US | 2003180856 |
| US | 2003180857 | US | 2003180858 | US | 2003180859 |
| US | 2003180860 | US | 2003180861 | US | 2003180862 |
| US | 2003180863 | US | 2003180867 | US | 2003180904 |
| US | 2003180905 | US | 2003180906 | US | 2003180907 |

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
**PCT/AU03/01166**

|    |            |    |            |    |            |
|----|------------|----|------------|----|------------|
| US | 2003180908 | US | 2003180909 | US | 2003180910 |
| US | 2003180911 | US | 2003180912 | US | 2003180913 |
| US | 2003180914 | US | 2003180915 | US | 2003180916 |
| US | 2003180917 | US | 2003180918 | US | 2003180919 |
| US | 2003180920 | US | 2003180921 | US | 2003180922 |
| US | 2003180923 | US | 2003181636 | US | 2003181637 |
| US | 2003181638 | US | 2003181639 | US | 2003181640 |
| US | 2003181641 | US | 2003181642 | US | 2003181643 |
| US | 2003181644 | US | 2003181645 | US | 2003181646 |
| US | 2003181647 | US | 2003181648 | US | 2003181649 |
| US | 2003181650 | US | 2003181651 | US | 2003181652 |
| US | 2003181653 | US | 2003181654 | US | 2003181655 |
| US | 2003181656 | US | 2003181657 | US | 2003181666 |
| US | 2003181667 | US | 2003181668 | US | 2003181669 |
| US | 2003181670 | US | 2003181671 | US | 2003181672 |
| US | 2003181673 | US | 2003181674 | US | 2003181675 |
| US | 2003181676 | US | 2003181677 | US | 2003181678 |
| US | 2003181679 | US | 2003181680 | US | 2003181681 |
| US | 2003181682 | US | 2003181683 | US | 2003181684 |
| US | 2003181685 | US | 2003181696 | US | 2003181697 |
| US | 2003181698 | US | 2003181699 | US | 2003181700 |
| US | 2003181701 | US | 2003181702 | US | 2003181703 |
| US | 2003185841 | US | 2003186318 | US | 2003186319 |
| US | 2003186365 | US | 2003186368 | US | 2003186407 |
| US | 2003186408 | US | 2003187189 | US | 2003187190 |
| US | 2003187191 | US | 2003187192 | US | 2003187193 |
| US | 2003187195 | US | 2003187196 | US | 2003187197 |
| US | 2003187228 | US | 2003187229 | US | 2003187230 |
| US | 2003187239 | US | 2003187241 | US | 2003187242 |
| US | 2003190321 | US | 2003190667 | US | 2003190668 |
| US | 2003190669 | US | 2003190698 | US | 2003190699 |
| US | 2003190701 | US | 2003190703 | US | 2003190716 |
| US | 2003190717 | US | 2003190718 | US | 2003190719 |
| US | 2003190720 | US | 2003190721 | US | 2003190722 |
| US | 2003190723 | US | 2003190724 | US | 2003190725 |
| US | 2003190726 | US | 2003190727 | US | 2003190728 |
| US | 2003190729 | US | 2003190730 | US | 2003190731 |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/AU03/01166

|    |            |    |            |    |            |
|----|------------|----|------------|----|------------|
| US | 2003191281 | US | 2003191282 | US | 2003191283 |
| US | 2003191284 | US | 2003191287 | US | 2003191288 |
| US | 2003191289 | US | 2003191290 | US | 2003191299 |
| US | 2003194744 | US | 2003194780 | US | 2003194781 |
| US | 2003194791 | US | 2003194792 | US | 2003194793 |
| US | 2003194794 | US | 2003195148 | US | 2003195333 |
| US | 2003195334 | US | 2003195344 | US | 2003195345 |
| US | 2003195347 | WO | 0032221    | WO | 0037638    |
| WO | 0037640    | WO | 0053752    | WO | 0053753    |
| WO | 0053755    | WO | 0053756    | WO | 0053757    |
| WO | 0053758    |    |            |    |            |
| WO | 0170976    | AU | 49317/01   | US | 2002048759 |
| US | 2003059781 |    |            |    |            |
| WO | 03068054   |    |            |    |            |
| US | 2002192678 |    |            |    |            |
| WO | 0192581    | AU | 65296/01   | CA | 2411404    |
|    |            | US | 2002132237 | EP | 1356092    |
| WO | 0151513    | AU | 36470/01   |    |            |
| WO | 0190304    | AU | 29508/01   | AU | 30958/01   |
|    |            | AU | 36460/01   | AU | 36459/01   |
|    |            | AU | 36461/01   | AU | 36462/01   |
|    |            | AU | 36463/01   | AU | 36465/01   |
|    |            | AU | 36466/01   | AU | 36466/01   |
|    |            | AU | 37943/01   | AU | 37944/01   |
|    |            | AU | 37947/01   | AU | 37949/01   |
|    |            | AU | 37951/01   | AU | 37950/01   |
|    |            | AU | 37952/01   | AU | 37953/01   |
|    |            | AU | 37954/01   | AU | 37955/01   |
|    |            | AU | 37955/01   | AU | 37957/01   |
|    |            | AU | 37958/01   | AU | 38585/01   |
|    |            | AU | 39727/01   | AU | 39726/01   |
|    |            | AU | 39728/01   | AU | 39728/01   |
|    |            | AU | 41403/01   | AU | 41402/01   |
|    |            | AU | 41404/01   | AU | 41405/01   |
|    |            | AU | 41406/01   | AU | 41407/01   |
|    |            | AU | 41407/01   | AU | 41408/01   |
|    |            | AU | 41409/01   | AU | 41410/01   |
|    |            | AU | 41410/01   | AU | 41411/01   |
|    |            | AU | 41412/01   | AU | 41413/01   |
|    |            | AU | 41413/01   | AU | 41414/01   |
|    |            | AU | 41415/01   | AU | 41416/01   |
|    |            | AU | 41416/01   | AU | 41417/01   |
|    |            | AU | 41418/01   | AU | 41419/01   |
|    |            | AU | 41419/01   | AU | 43134/01   |
|    |            | AU | 43135/01   | AU | 43136/01   |
|    |            | AU | 43136/01   | AU | 43137/01   |
|    |            | AU | 45262/01   | AU | 45354/01   |
|    |            | AU | 45354/01   | AU | 47190/01   |
|    |            | AU | 47191/01   | AU | 49052/01   |
|    |            | AU | 49052/01   | AU | 49053/01   |
|    |            | AU | 49054/01   | AU | 50767/01   |
|    |            | AU | 50767/01   | AU | 50768/01   |
|    |            | AU | 50769/01   | AU | 50770/01   |
|    |            | AU | 50770/01   | AU | 50771/01   |



**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
**PCT/AU03/01166**

|             |             |             |
|-------------|-------------|-------------|
| AU 50772/01 | AU 52878/01 | AU 52879/01 |
| AU 55162/01 | AU 57912/01 | AU 60969/01 |
| AU 62899/01 | AU 74888/01 | CA 2392398  |
| CA 2392422  | CA 2392428  | CA 2392438  |
| CA 2392450  | CA 2392751  | CA 2392757  |
| CA 2393002  | CA 2393616  | CA 2393618  |
| CA 2393652  | CA 2393912  | CA 2393941  |
| CA 2393954  | CA 2394022  | CA 2394039  |
| CA 2394841  | CA 2395178  | CA 2395295  |
| CA 2395398  | CA 2395403  | CA 2395654  |
| CA 2395666  | CA 2395671  | CA 2395676  |
| CA 2395693  | CA 2395699  | CA 2395724  |
| CA 2395729  | CA 2395734  | CA 2395738  |
| CA 2395787  | CA 2395794  | CA 2395811  |
| CA 2395815  | CA 2395816  | CA 2395827  |
| CA 2395838  | CA 2395849  | CA 2395857  |
| CA 2395858  | CA 2395872  | CA 2395885  |
| CA 2395889  | CA 2396719  | CA 2397407  |
| CA 2397502  | CA 2397839  | CA 2398227  |
| CA 2398275  | CA 2398411  | CA 2398877  |
| EP 1251863  | EP 1252176  | EP 1252185  |
| EP 1252289  | EP 1252290  | EP 1252297  |
| EP 1252302  | EP 1252303  | EP 1252312  |
| EP 1252326  | EP 1252337  | EP 1254147  |
| EP 1254148  | EP 1254150  | EP 1254151  |
| EP 1254152  | EP 1254153  | EP 1254154  |
| EP 1254157  | EP 1254165  | EP 1254170  |
| EP 1254171  | EP 1254172  | EP 1254173  |
| EP 1254217  | EP 1254218  | EP 1254219  |
| EP 1254228  | EP 1254242  | EP 1254248  |
| EP 1254272  | EP 1255766  | EP 1255767  |
| EP 1255768  | EP 1255776  | EP 1255777  |
| EP 1255778  | EP 1255817  | EP 1255864  |
| EP 1255869  | EP 1259525  | EP 1259526  |
| EP 1259528  | EP 1259531  | EP 1259540  |
| EP 1259642  | EP 1261380  | EP 1261618  |
| EP 1261619  | EP 1261633  | EP 1261634  |

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
**PCT/AU03/01166**

|    |            |    |            |    |            |
|----|------------|----|------------|----|------------|
| EP | 1261635    | EP | 1261637    | EP | 1261703    |
| EP | 1261742    | EP | 1261745    | EP | 1263944    |
| EP | 1265910    | EP | 1268527    | US | 2002042096 |
| US | 2002042386 | US | 2002061521 | US | 2002077270 |
| US | 2002086330 | US | 2002086353 | US | 2002086811 |
| US | 2002086820 | US | 2002086821 | US | 2002086822 |
| US | 2002086823 | US | 2002090615 | US | 2002090672 |
| US | 2002090673 | US | 2002090674 | US | 2002094953 |
| US | 2002102638 | US | 2002119919 | US | 2002120103 |
| US | 2002132753 | US | 2002132767 | US | 2002147140 |
| US | 2002151009 | US | 2002151479 | US | 2002160493 |
| US | 2002161208 | US | 2002164685 | US | 2002165137 |
| US | 2002168711 | US | 2002173454 | US | 2003013649 |
| US | 2003039993 | US | 2003039994 | US | 2003044890 |
| US | 2003044904 | US | 2003044905 | US | 2003044907 |
| US | 2003049650 | US | 2003049652 | US | 2003049703 |
| US | 2003050231 | US | 2003054368 | US | 2003054373 |
| US | 2003054375 | US | 2003054377 | US | 2003054379 |
| US | 2003054420 | US | 2003059875 | US | 2003059908 |
| US | 2003077602 | US | 2003077606 | US | 2003077703 |
| US | 2003077704 | US | 2003077808 | US | 2003082681 |
| US | 2003082758 | US | 2003092102 | US | 2003092611 |
| US | 2003092615 | US | 2003096346 | US | 2003108907 |
| US | 2003125246 | US | 2003139327 | US | 2003157508 |
| US | 2003175739 | US | 2003190707 | WO | 0154472    |
| WO | 0154473    | WO | 0154474    | WO | 0154708    |
| WO | 0154733    | WO | 0155162    | WO | 0155163    |
| WO | 0155164    | WO | 0155167    | WO | 0155168    |
| WO | 0155173    | WO | 0155200    | WO | 0155201    |
| WO | 0155202    | WO | 0155203    | WO | 0155204    |
| WO | 0155205    | WO | 0155206    | WO | 0155207    |
| WO | 0155208    | WO | 0155300    | WO | 0155301    |
| WO | 0155302    | WO | 0155303    | WO | 0155304    |
| WO | 0155305    | WO | 0155306    | WO | 0155307    |
| WO | 0155308    | WO | 0155309    | WO | 0155310    |
| WO | 0155311    | WO | 0155312    | WO | 0155313    |
| WO | 0155314    | WO | 0155315    | WO | 0155316    |

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
**PCT/AU03/01166**

|            |               |               |
|------------|---------------|---------------|
| WO 0155317 | WO 0155318    | WO 0155319    |
| WO 0155320 | WO 0155321    | WO 0155322    |
| WO 0155323 | WO 0155324    | WO 0155325    |
| WO 0155326 | WO 0155327    | WO 0155328    |
| WO 0155329 | WO 0155343    | WO 0155350    |
| WO 0155355 | WO 0155364    | WO 0155367    |
| WO 0155368 | WO 0155387    | WO 0155388    |
| WO 0155430 | WO 0155440    | WO 0155441    |
| WO 0155447 | WO 0155448    | WO 0155449    |
| WO 0157182 | WO 0159063    | WO 0159064    |
| WO 0162789 | WO 0164703    |               |
| WO 9953040 | DE 19817557   | EP 1073727    |
|            | US 2003105315 | EP 1291422    |
| WO 0194629 | AU 64559/01   | CA 2411601    |
|            | US 2002102531 | US 2002081301 |
|            | US 2002115057 | US 2002110821 |
|            | US 2002165180 | US 2002150877 |
|            | US 2003165839 |               |
| WO 9848010 | AU 70139/98   | EP 0979277    |
|            | US 2002192160 |               |
| WO 0160860 | AU 41541/01   |               |
| WO 0155313 | AU 29508/01   | AU 30958/01   |
|            | AU 36459/01   | AU 36460/01   |
|            | AU 36462/01   | AU 36463/01   |
|            | AU 36465/01   | AU 36466/01   |
|            | AU 37944/01   | AU 37943/01   |
|            | AU 37947/01   | AU 37949/01   |
|            | AU 37950/01   | AU 37951/01   |
|            | AU 37952/01   | AU 37953/01   |
|            | AU 37954/01   | AU 37955/01   |
|            | AU 37957/01   | AU 37958/01   |
|            | AU 38585/01   | AU 39726/01   |
|            | AU 39727/01   | AU 39728/01   |
|            | AU 41402/01   | AU 41403/01   |
|            | AU 41404/01   | AU 41405/01   |
|            | AU 41406/01   | AU 41407/01   |
|            | AU 41408/01   | AU 41409/01   |
|            | AU 41410/01   | AU 41411/01   |
|            | AU 41412/01   | AU 41413/01   |
|            | AU 41414/01   | AU 41415/01   |
|            | AU 41416/01   | AU 41417/01   |
|            | AU 41418/01   | AU 41419/01   |
|            | AU 43134/01   | AU 43135/01   |
|            | AU 43136/01   | AU 43137/01   |
|            | AU 45262/01   | AU 45354/01   |
|            | AU 47190/01   | AU 47191/01   |
|            | AU 49052/01   |               |

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
**PCT/AU03/01166**

|             |             |             |
|-------------|-------------|-------------|
| AU 49053/01 | AU 49054/01 | AU 50767/01 |
| AU 50768/01 | AU 50769/01 | AU 50770/01 |
| AU 50771/01 | AU 50772/01 | AU 52878/01 |
| AU 52879/01 | AU 52917/01 | AU 53323/01 |
| AU 55162/01 | AU 57912/01 | AU 60969/01 |
| AU 62899/01 | AU 66787/01 | AU 71714/01 |
| AU 74888/01 | CA 2392398  | CA 2392422  |
| CA 2392428  | CA 2392438  | CA 2392450  |
| CA 2392751  | CA 2392757  | CA 2393002  |
| CA 2393616  | CA 2393618  | CA 2393652  |
| CA 2393912  | CA 2393941  | CA 2393954  |
| CA 2394022  | CA 2394039  | CA 2394841  |
| CA 2395178  | CA 2395295  | CA 2395398  |
| CA 2395403  | CA 2395654  | CA 2395666  |
| CA 2395671  | CA 2395676  | CA 2395693  |
| CA 2395699  | CA 2395724  | CA 2395729  |
| CA 2395734  | CA 2395738  | CA 2395787  |
| CA 2395794  | CA 2395811  | CA 2395815  |
| CA 2395816  | CA 2395827  | CA 2395838  |
| CA 2395849  | CA 2395857  | CA 2395858  |
| CA 2395872  | CA 2395885  | CA 2395889  |
| CA 2396719  | CA 2397407  | CA 2397502  |
| CA 2397839  | CA 2398227  | CA 2398275  |
| CA 2398411  | CA 2398877  | CA 2400638  |
| CA 2403508  | CA 2406649  | EP 1251863  |
| EP 1252176  | EP 1252185  | EP 1252289  |
| EP 1252290  | EP 1252297  | EP 1252302  |
| EP 1252303  | EP 1252312  | EP 1252326  |
| EP 1252337  | EP 1254147  | EP 1254148  |
| EP 1254150  | EP 1254151  | EP 1254152  |
| EP 1254153  | EP 1254154  | EP 1254157  |
| EP 1254165  | EP 1254170  | EP 1254171  |
| EP 1254172  | EP 1254173  | EP 1254217  |
| EP 1254218  | EP 1254219  | EP 1254228  |
| EP 1254242  | EP 1254248  | EP 1254272  |
| EP 1255766  | EP 1255767  | EP 1255768  |
| EP 1255776  | EP 1255777  | EP 1255778  |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/AU03/01166

|    |            |    |            |    |            |
|----|------------|----|------------|----|------------|
| EP | 1255817    | EP | 1255864    | EP | 1255869    |
| EP | 1259525    | EP | 1259526    | EP | 1259528    |
| EP | 1259531    | EP | 1259540    | EP | 1259642    |
| EP | 1261380    | EP | 1261618    | EP | 1261619    |
| EP | 1261633    | EP | 1261634    | EP | 1261635    |
| EP | 1261637    | EP | 1261703    | EP | 1261742    |
| EP | 1261745    | EP | 1263944    | EP | 1265910    |
| EP | 1268527    | EP | 1268541    | EP | 1287015    |
| EP | 1301524    | US | 2002042096 | US | 2002042386 |
| US | 2002061521 | US | 2002077270 | US | 2002086330 |
| US | 2002086353 | US | 2002086811 | US | 2002086820 |
| US | 2002086821 | US | 2002086822 | US | 2002086823 |
| US | 2002090615 | US | 2002090672 | US | 2002090673 |
| US | 2002090674 | US | 2002094953 | US | 2002102638 |
| US | 2002106780 | US | 2002119919 | US | 2002120103 |
| US | 2002132753 | US | 2002132767 | US | 2002147140 |
| US | 2002151009 | US | 2002151479 | US | 2002160493 |
| US | 2002161208 | US | 2002164685 | US | 2002165137 |
| US | 2002168711 | US | 2002173454 | US | 2003013649 |
| US | 2003039993 | US | 2003039994 | US | 2003044890 |
| US | 2003044904 | US | 2003044905 | US | 2003044907 |
| US | 2003049650 | US | 2003049652 | US | 2003049703 |
| US | 2003050231 | US | 2003054368 | US | 2003054373 |
| US | 2003054375 | US | 2003054377 | US | 2003054379 |
| US | 2003054420 | US | 2003059875 | US | 2003059908 |
| US | 2003077602 | US | 2003077606 | US | 2003077703 |
| US | 2003077704 | US | 2003077808 | US | 2003082681 |
| US | 2003082758 | US | 2003092102 | US | 2003092611 |
| US | 2003092615 | US | 2003096346 | US | 2003108907 |
| US | 2003125246 | US | 2003139327 | US | 2003157508 |
| US | 2003175739 | US | 2003190707 | WO | 0154472    |
| WO | 0154473    | WO | 0154474    | WO | 0154708    |
| WO | 0154733    | WO | 0155162    | WO | 0155163    |
| WO | 0155164    | WO | 0155167    | WO | 0155168    |
| WO | 0155173    | WO | 0155200    | WO | 0155201    |
| WO | 0155202    | WO | 0155203    | WO | 0155204    |
| WO | 0155205    | WO | 0155206    | WO | 0155207    |

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☒ **BLACK BORDERS**

☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

☐ **FADED TEXT OR DRAWING**

☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**

☐ **SKEWED/SLANTED IMAGES**

☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**

☐ **GRAY SCALE DOCUMENTS**

☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**

☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**